Development of a Flexible Framework for Deterioration Modelling in Infrastructure Asset Management

by

Abra Ens

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Department of Civil Engineering
University of Toronto

A Framework for Deterioration Modelling Development in Infrastructure Asset Management

Abra Ens

Master of Applied Science

Department of Civil Engineering
University of Toronto

2012

Abstract

Infrastructure deterioration models are an integral part of asset management. Deterioration models are used to predict future asset condition and to estimate funding requirements.

The purpose of this research is to develop a framework to create infrastructure deterioration models. An overview of the various types of deterioration models is included, presenting the advantages and disadvantages of each type. Existing deterioration model frameworks are also considered. A deterioration modelling framework is then proposed. The selection of the model type, calibration and validation is presented.

The framework is then applied to two case studies. The first case study involves a comparison of three pavement deterioration models, created for the City of Oshawa for use in their asset management system. The second case study involves modelling sewer deterioration. This model has been developed to explore the relationship between age, material and deterioration in trunk sewers.

Table of Contents

List of Tables

List of Figures

1 Introduction

According to the International Infrastructure Management Manual (2006), “Infrastructure assets are stationary systems (or networks) that serve defined communities where the system as a whole is intended to be maintained indefinitely to a specified level of service by continuing replacement and refurbishment of its components.” Thus, the term “infrastructure” encompasses a wide breadth of assets, including roads, sewers, parks, public buildings and telecommunication networks, to name a few. Municipal infrastructure can be classified as linear (e.g. roads, sewers, etc.), or point (e.g. bridges, schools, water treatment plants, etc.).

In the late 1990s, over $100 billion was spent each year on maintenance, repair and capital renewal of municipal infrastructure in Canada (Vanier 2001). It must be assumed that the quantity spent on these activities has only been increasing as populations grow and infrastructure ages. To effectively manage municipal infrastructure networks, asset management plans are developed. The goal of these plans is to, “meet a required level of service, in the most cost effective manner, through the management of assets for present and future customers” (IIMM 2006). Asset management is particularly important in Canada because its infrastructure, most of which was built in the 1950’s and 1960’s, is aging and failing, sometimes in catastrophic ways. Thus, the maintenance and rehabilitation of existing municipal infrastructure is becoming increasingly important.

Infrastructure asset management involves creating an asset inventory with records for all assets in the system; assessing the condition of these assets, typically through inspection; and finally analyzing this data in conjunction with budget information and required levels of service. From this, an overall maintenance and rehabilitation schedule is created which is also used to predict future funding needs.

Asset management systems typically operate at two levels: network and project. At the network level, the optimal maintenance and rehabilitation schedule is found subject to budget and level of service constraints. At the project level, decisions are made involving which maintenance and rehabilitation technique to use on a particular asset.

Deterioration models are used at both levels to estimate an asset's condition. At the network level, condition is typically measured as either a functional or structural measure and converted to an aggregated, index type variable. At the project level, condition may be measured at a finer scale, perhaps estimating the severity and extent of each distress individually, so that appropriate treatments can be recommended. For example, pavement deterioration could be measured at the network level using the International Roughness Index (IRI) and at the project level as a pavement distress.

1.1 Objectives

The purpose of this research is to develop a framework to create infrastructure deterioration models. The framework should be flexible and comprehensive, and result in a model that can easily be incorporated into an asset management system.

1.2 Scope

Due to space constraints, it is not possible to provide an in-depth discussion on deterioration mechanisms, and existing models for multiple types of municipal infrastructure. This thesis focuses on linear municipal assets, particularly roads and sewers. However, the proposed framework can be applied to other infrastructure assets as well.

This thesis presents a literature review outlining the deterioration process and the influencing factors for roads and sewers. An overview of the types of deterioration models is also included, presenting both the advantages and disadvantages of each. Existing deterioration model frameworks are also considered. A deterioration modelling framework is then proposed and evaluated using real world data. The selection of the deterioration model, calibration and validation is presented.

The framework is applied to two case studies. The first is a comparison of three pavement deterioration models created for the City of Oshawa for use in their asset management system. The second case study involves modelling sewer deterioration, where the relationship between age, material and deterioration in large trunk sewers is explored.

2 Literature Review

2.1 Deterioration Process

Deterioration is a function of by environment loading, structural loading, and various other factors. External factors include the number of freeze/thaw cycles, traffic loading (for pavements) and type of waste transported (for sewers). Intrinsic factors include material type and construction. Maintenance factors include the type and frequency of maintenance treatments.

In many infrastructure assets, the rate of deterioration is expected to gradually increase with time. A typical deterioration curve (without maintenance activities) can be found in Figure 1.

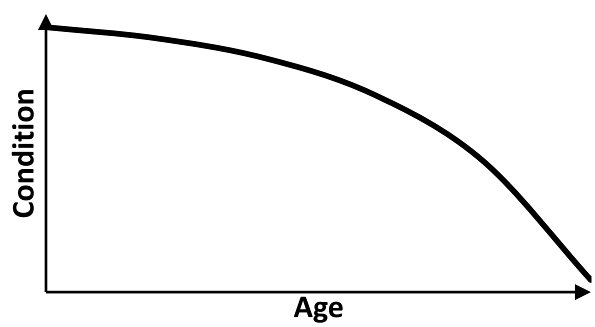


Figure 1 Typical deterioration curve (NGSMI 2002)

However, deterioration does not always occur in this way. Concave up deterioration curves are found when pavements have been designed to a higher standard than required for traffic alone, and primarily deteriorate due to weather/climate factors (Haas 1997). Also, a single damage event may cause an asset to deteriorate very rapidly or almost instantaneously.

Another way of evaluating deterioration is through the probability of failure. The "bathtub curve" (Figure 2) is used in reliability engineering and has been applied to pipes (Kleiner et al. 2001).

Figure 2 "Bathtub" curve (Kleiner et al. 2001)

In the first section, with a decreasing failure rate, the asset is prone to failure due to “infant mortality” – typically due to construction or fabrication errors. In the second section, assets may experience random failure, but typically have a low failure rate. In the third section, where the failure rate begins to increase again, the assets fail due to age and use – they become worn out.

Asset deterioration models – particularly pavement deterioration models – are used not only in the management of infrastructure, but also in design (C-SHRP, 2000) and performance specifications for contractors (Parkman et al. 2003). As the models for other assets become more accurate and reliable, their use might be extended as well.

2.2 Factors that Affect Deterioration

To create a deterioration model, the factors that affect the infrastructure’s condition must be quantified. For example, the causes of pavement deterioration are well-known, and include environmental, traffic and structural factors. Environmental factors can include measurements of the number of freeze-thaw cycles, temperature, humidity, precipitation, water table depth; traffic factors typically include measurements of Average Annual Daily Traffic (AADT) or Equivalent Single Axle Loads (ESALs). Structural factors can include pavement type, strength and thickness, Granular Base Equivalency (GBE), subgrade material, and existing pavement distress measurements.

Construction and maintenance techniques also influence pavement deterioration. However these factors can be more difficult to quantify and are not included in many deterioration models. Location factors are an important consideration when modelling pavement deterioration (Raymond et al. 2003). Some international models, including the World Bank's Highway Development and Management Model HDM-4 (Kerali et al. 1998), suggest that the user calibrate the model to account for regional differences.

Although the general deterioration process for storm water and sanitary sewers is well documented, there is no consensus as to which factors influence their rate of deterioration. Table 1 provides an example of five (5) factors and whether the factor was found to be significant in a selection of studies.

Table 1 Significance of independent variables in literature for sewer models

Factor	(Ariaratnam et al. 2001)	(Davies 2001)	(Baik et al. 2006)	(Tran et al. 2009)	(Ana et al. 2009)	Number of Significant Instances
Age	sig	not	sig	not	sig	3 of 5
Depth	not	not	no data	not	not	0 of 4
Location	no data	sig	no data	sig	not	2 of 3
Material	not	sig	not	no data	sig	2 of 4
Size	sig	sig	sig	sig	not	4 of 5

*sig =significant, not=not significant

Therefore, it cannot be assumed that any factor is significant in a given situation. Even a factor such as age, which appears to be inherently linked to deterioration, was not significant in several studies. There are several possible reasons for the differences between the studies shown in Table 1 (Scheidegger et al. 2011):

- Because sewer deterioration is an inherently complex process, with many variables interacting and affecting one another, it is extremely unlikely that any one data set will capture all of the variables and interactions. The lack of availability of data means that each model starts with a different combination of variables and that between studies, these variables may not be comparable. For example, one model might include only rigid pipes (e.g. concrete, vitrified clay, etc.) in the "material" category, while another might also include flexible pipes (e.g. high-density polyethylene, polyvinyl chloride, etc.). If the difference between flexible and rigid pipes is significant, but the difference between materials within the rigid pipe category (e.g. the difference between concrete and vitrified clay) is not, one study may conclude that "material" is a significant factor while the other will not.
- Pipe networks have been constructed and repaired over many years using different methods, materials and specifications. By categorizing pipes into discrete groups, a complicated problem may be oversimplified.
- The statistical selection of variables may eliminate those that are correlated. For example, pipe size generally increases with depth so one of these variables will typically be eliminated from the model.

2.3 Types of Deterioration Models

Deterioration models can be classified as deterministic and probabilistic. In addition, various other techniques such as artificial intelligence can be used to develop performance models. The following section will present an overview of each group of models, along with a more in-depth review of an example of each type.

2.3.1 Deterministic Models

A deterministic model outputs a single condition value for a given set of inputs. Deterministic models are typically displayed as functions. The simplest of these types of models is created using linear regression, but exponential and other, more complex functions can produce more accurate results.

Deterministic models are either mechanistic, empirical, mechanistic-empirical or based on expert opinion. Mechanistic models are based on physical laws. For example, in

infrastructure deterioration modelling, the relationships between stress, strain, loading and deflection may be used. Mechanistic models are not typically used in asset deterioration modelling because deterioration is usually caused by the interaction of many different factors which mechanistic models cannot account for.

Empirical models are developed by relating condition scores to explanatory variables (such as age, material type, loading conditions, etc) usually through a regression process. This type of model is frequently employed when deterioration cannot be explained by mechanic processes. Schram (2008) found that 91% of the responding Canadian and American agencies used empirical pavement deterioration models. Most Nordic countries also use an empirical linear extrapolation of a pavement's current condition in their asset management systems (Saba 2006). Although not frequently used to model pipe deterioration, a deterministic empirical model has been created to model sewer deterioration caused by corrosion (Konig 2005).

Many deterioration models fall into the category of mechanistic-empirical. Mechanistic-empirical models incorporate calculated mechanistic responses (e.g. strain, deformation, etc.) and other measured variables to predict condition. This type of model is frequently used to model pavement deterioration (Raymond et al. 2003; Schram 2008; Ullidtz 1999; Tighe et al. 2001) has been shown to give good results, and is thought to model deterioration more accurately than empirical models alone. Mechanistic empirical models have also been used to model water main deterioration (Rajani et al. 2001).

2.3.1.1 Multiple Linear Regression

Multiple linear regression is one of the simplest forms of a deterministic model and is used when more than one factor influence the dependent variable. The model is estimated to fit the equation:

$$\hat{y}=b_0+b_1x_1+\ldots+b_kx_k \qquad [1]$$

where b_0, $b_1,\ldots,b_k$ are the estimates of the regression coefficients, $\hat{y}$ is the predicted value of the dependent variable, and $x_1,\ldots, x_k$ are the values of the independent variables. In the case of infrastructure deterioration, $\hat{y}$ is generally the condition of the asset, and $x_1,\ldots, x_k$ are the

factors that affect the asset's condition (e.g. age, material, location, etc…). To find the value of the coefficients, $b_0, \ldots, b_k$, the method of least squares is commonly used.

One difficulty that may be encountered when attempting to use multiple linear regression to model asset deterioration is that the condition value in condition assessment surveys is often measured on a discrete scale (such as the Water Research Council (WRc) system commonly used for sewers) rather than a continuous one, as is assumed in multiple linear regression.

2.3.2 Probabilistic Models

Probabilistic models, on the other hand, output a probability that an asset is in a particular condition given a set of inputs. Probabilistic models are frequently used to model infrastructure deterioration and many different model types exist within this group.

2.3.2.1 Markov Models

One of the most popular probabilistic models used to model asset deterioration is the Markov chain. Markov models give the probability, p_{ij}, that an element in state i at time-step t, will be in state j at time-step (t+1). These transition probabilities are assembled in the form of a transition matrix.

$$P^{t,t+1}=P(X_{t+1}=j|X_t=i)=\begin{bmatrix} p_{11} & \cdots & p_{1j} \\ \vdots & \ddots & \vdots \\ p_{i1} & \cdots & p_{ij} \end{bmatrix} \qquad [2]$$

where $p_{ij} \geq 0$; $i,j \geq 1$; $\sum_{k=1}^{j} p_{i,k}=1$.

The distribution of the states of an asset network at time (t+n) can be found by taking the product of the current distribution and the transition matrices:

$$Q(t+n)=Q(t)P^{t,t+1}P^{t+1,t+2}\ldots P^{t+n-1,t+n} \qquad [3]$$

When modelling infrastructure deterioration, p_{ij} is usually defined as the probability of an asset deteriorating from condition i to condition j. When $i > j$, $P_{ij} = 0$ unless rehabilitation or repair has taken place.

While in a time-homogeneous Markov model (shown above), future states are dependent only on the present state, in a semi-Markov, or non-homogeneous model, independently distributed random variables are used to model the time between the states. Thus, the model is time-dependent. In terms of asset deterioration, this means that the probability of deteriorating to the next state increases with the age of the asset. The semi-Markov model requires more data for its extra parameters and has a more complex implementation than a time-homogeneous Markov (Black et al. 2005).

Typically, the most challenging aspect to creating a Markov model is to determine the probabilities in the transition matrix. If the model is semi-Markov, there is the additional challenge of modelling the time between states.

Transitional probabilities in a Markov deterioration model can be estimated using many methods, including Weibull distribution (Kleiner 2001), non-linear optimization to fit an exponential regression model to historical data (Wirahadikusumah et al. 2001), Bayesian inference in combination with the Metropolis-Hastings algorithm (Micevski et al. 2002), an ordered probit model (Baik et al. 2006), and the Gompit model, an extension of the probit model (Le Gat 2008). Expert opinion can also be used to estimate the values of the parameters needed for the model if sufficient historical data is not available (Kleiner 2001).

2.3.2.2 Probabilistic Regression Models

Another commonly used probabilistic model is logistic regression. Unlike multiple linear regression, where the output is the condition state of the asset; logistic regression provides the probability that an asset is in a particular condition state given a set of independent variables. The probability is written in terms of the logistic function:

$$E(Y=y|X)=P=\frac{1}{1+e^{-(b_0+\sum_{i=1}^{k} b_i x_i)}} \qquad [4]$$

where, with respect to asset deterioration modelling, P is the probability that the asset is in a particular discrete state, or condition, y; x_i are the factors that affect an asset's condition, and b_i are the estimates of the regression coefficients. It follows, then, that since Y is binary, the probability that the asset is not in condition y is $(1-P)$. To find the value of the regression

coefficients, the maximum likelihood method is commonly used. The logistic regression model can be extended to include more than one category for the dependent variable by using ordinal logistic regression.

The probit model is very similar to the logistic regression model. The difference between the two is found in the underlying distribution function: in a logistic regression model, the distribution is logistic, whereas in a probit model, the distribution is standard normal. This causes the logistic model to have flatter tails than the probit model. Both models produce similar results, but the logistic model has two advantages over the probit model. The logistic model is computationally simpler, and the odds ratio that can be found using logistic regression is easily understood.

Successful applications of probabilistic regression models for asset deterioration have been developed in Canada (Younis et al. 2010b) and internationally (Davies 2001; Ana et al. 2009; Henning et al. 2006).

2.3.2.3 Other Probabilistic Models

Other probabilistic methods used to model infrastructure deterioration include multiple discriminant analysis (Tran et al. 2006), cohort survival (Baur et al. 2002) and proportional hazards models (Yu 2005).

2.3.3 Artificial Intelligence Methods

Soft computing methods are often modelled on processes found in nature, such as the brain, or natural selection. Soft computing techniques allow for uncertain, imprecise, and ambiguous data. Because this often describes asset inventories and condition information, soft computing methods have been used to create infrastructure deterioration models (Flintsch et al. 2004).

2.3.3.1 Artificial Neural Networks

A popular artificial intelligence method is the artificial neural network. Artificial neural network (ANN) models are based on the structure of the neurons in a brain, where each neuron processes its inputs, and then outputs a signal, or value, which is passed on to the next

neuron. With a number of such neurons working in parallel, a final output can then be determined. Figure 3 illustrates the structure of an artificial neural network.

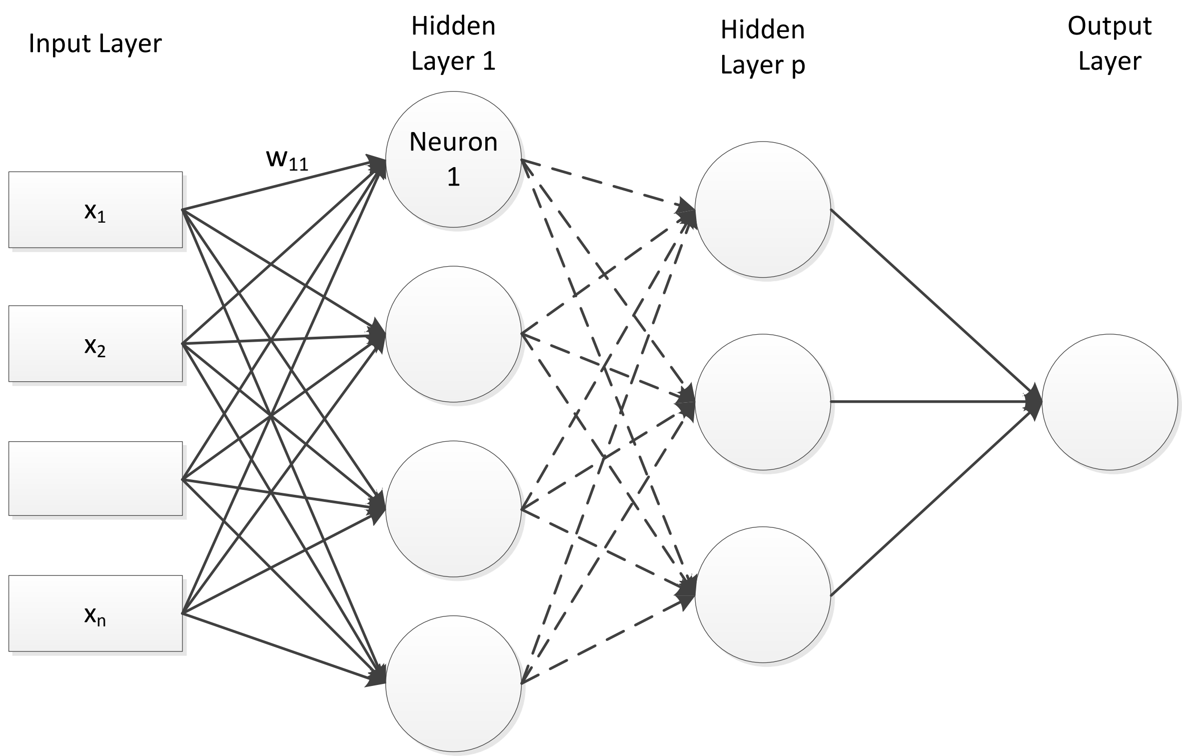


Figure 3 Structure of an artificial neural network

Like regression models, ANN models begin with the values of n factors, x_i. Each of these input variables is multiplied by a weight, w_{ik}, and is summed in neuron k to find an activation value, a_k, for that particular neuron.

$$a_k = \sum_{i=1}^{n} w_{ik} x_i \qquad [5]$$

where n is the number of neurons or inputs connected to neuron k, a_k is the activation value for neuron k, w_{ik} is the weight associated with input i of neuron k, and x_i is the value of input i.

Then, the activation value is transformed, using a sigmoid function (a mathematical function having an "S" shape), into an output value, O_k, usually between 0 and 1. For example, the logistic function may be used:

$$O_k = \frac{1}{1+e^{-g_k a_k}} \quad [6]$$

where O_k is the output of neuron k, a_k is the activation value of neuron k, and g_k is the gain of the sigmoid transfer function of neuron k.

This output, O_k , then becomes an input variable for the next layer of neurons, which continues until a final output for the system is reached.

There are many ways to train an ANN (i.e. to determine the values of w_{ik}), but generally, input variables for which the desired output is known, are fed into the network and the weights for the connections between neurons are changed until the squared difference between the actual outcomes and the desired outcomes are acceptable. This is repeated hundreds or thousands of times for each set of outputs and input variables in the training set.

Neural networks have been widely used to model performance for infrastructure assets, including pavements (Eldin et al. 1995; Fwa et al. 1993; Lou et al. 2001),water mains (Al-Barqawi et al. 2008; Bubtiena et al. 2011), bridges (Cattan et al. 1997; Elhag et al. 2007), sanitary sewers (Najafi et al. 2005) and storm-water pipes (Tran et al. 2007). Many of these models aim to identify distressed segments (rather than the network condition as a whole) to prioritize inspections and predict maintenance needs.

2.3.3.2 Other Artificial Intelligence Models

Other soft computing methods commonly used in infrastructure deterioration modelling include genetic algorithms (Shekharan 2000; Chang et al. 2008), and fuzzy logic systems (Kleiner et al. 2006; Wang et al. 2011).

2.3.4 Advantages and Disadvantages of Model Types

The models described above have advantages and disadvantages when applied to infrastructure deterioration modelling. These are outlined in Table 2.

Table 2 Advantages and disadvantages of model types

Model Type	Examples	Advantages	Disadvantages
Deterministic Models	- multiple linear regression - exponential regression - non-linear regression	- provides insight into which factors most affect the deterioration process - final form (equation) is very user- friendly - relatively easy to understand and develop	- underlying assumptions, which can be difficult to validate, must be satisfied - not appropriate to model discrete states with a linear model (Madanat et al. 1995)

Probabilistic Models	- Markov models
	- probabilistic regression
	    - logistic
	    - probit
	- multiple discriminant analysis
	- cohort survival model
	- proportional hazard model
		- can be easily incorporated into risk models (Ana et al. 2010)
		- output discrete data (Tran 2007)
		- models the inherent uncertainty in the deterioration processes
			- may require longitudinal data that is not easily found (Baik et al. 2006; Wirahadikusumah et al. 2001; Madanat et al. 1995);
			- cohorts may need to be created (e.g.(Wirahadikusumah et al. 2001)), requiring more data

Artificial Intelligence Models	- artificial neural networks - genetic algorithms - fuzzy logic systems	- can model unknown, complex, nonlinear relationships between inputs and outputs - few underlying assumptions - can be used when data is imprecise, incomplete and subjective (Flintsch et al. 2004)	- more difficult to determine the significance of outputs (although they can be ranked (Olden et al. 2002)) - initial set-up can be time-consuming and complicated (Tran et al. 2010) - impossible to integrate prior knowledge for some training algorithms (Flintsch et al. 2004) - “black-box” technique means the path to the solution is not transparent - large amount of data needed for training and calibration (Scheidegger et al. 2011)

2.4 Deterioration Model Frameworks

Frameworks used to develop infrastructure deterioration models are not the subject of much literature. Typically, research focuses on the results of a particular model, rather than the process of its development.

Chughtai et al. (2008) presents a detailed framework that was used to create a non-linear regression model. While details on the model development process and the statistics used to evaluate the model are included, Chughtai's model is specific to non-linear regression models and cannot be used to develop other types of models. Similarly, Syachrani et al. (2011) focuses on adapting a typical regression model into a dynamic format. Like Chughtai's framework, this framework cannot be applied outside of its original model type.

Osman et al. (2011) provides a framework for the development of statistical deterioration models for water mains. Information on data cleansing and extraction, and details on the model selection process are explored in this research. However, this framework does not consider the iterative process of model development. In most cases, the first attempt to create a model will not be successful. Aspects of the model (be they the model form, independent variables, or even the model type) will likely need to be revised over the course of the model's development.

2.5 Summary of Findings

Infrastructure deterioration is a function of environmental loading, structural loading and various other factors. The factors that affect pavement deterioration are well-known, whereas there is no consensus on the factors that affect sewer deterioration. Deterioration can increase gradually with time, or occur in discrete steps.

Deterioration models can be classified as deterministic, probabilistic, or based on artificial intelligence. Deterministic models are easy to understand and develop, but generally assume gradual deterioration, which may not be valid for some asset types (Madanat et al. 1995). Probabilistic model output discrete data and quantify the inherent uncertainty of the deterioration process, but may require more data than is readily available (Wirahadikusumah et al. 2001). Artificial intelligence methods can model complex relationships between variables; but are time-consuming (Tran et al. 2010), require a large quantity of data (Scheidegger et al. 2011), and provide a "black-box" solution.

There are not many frameworks for infrastructure deterioration modelling found in literature. Frameworks that have been developed are particular to a certain type of model (Chughtai et

al. 2008; Syachrani et al. 2001) or do not consider the iterative process of model development (Osman et al. 2011).

3 Framework Development

The proposed deterioration modelling framework is presented in Figure 4.

3.1 Step 1a – Compile Data

The validity of a deterioration model is based on the accuracy and reliability of its data. This step entails taking several sources of data and combining them to create a comprehensive dataset. Depending on the data management techniques used to store existing data, this might be a time-consuming process. The key factors that might affect deterioration (based on experience and knowledge of the deterioration process) must be identified. Depending on the size of the available dataset (number of records and number of variables), and the existing location of the data, compilation of the data could be performed in an asset management software tool, in a database, or in a spreadsheet application.

To create a deterioration model, condition data and information relating to the characteristics of individual assets are necessary. Pavement condition data are relatively easy to collect. Roads are typically easy to access and several objective measurement methods and devices exist to evaluate condition. Buried infrastructure, on the other hand, is more difficult and time-consuming to access and has fewer methods of condition assessment. For example, sewer condition data, relative to other types of infrastructure, is difficult and expensive to collect and is generally subjective; the quality of the data depends on the skill of the inspector.

For deterioration models to be developed, a certain quantity of data is needed; generally, the accuracy of the model can be increased with a greater quantity of data. The problem of too few data can be exacerbated when models are divided into cohorts (for example, some Markov models (Ana et al. 2010) or are inherently data-hungry.

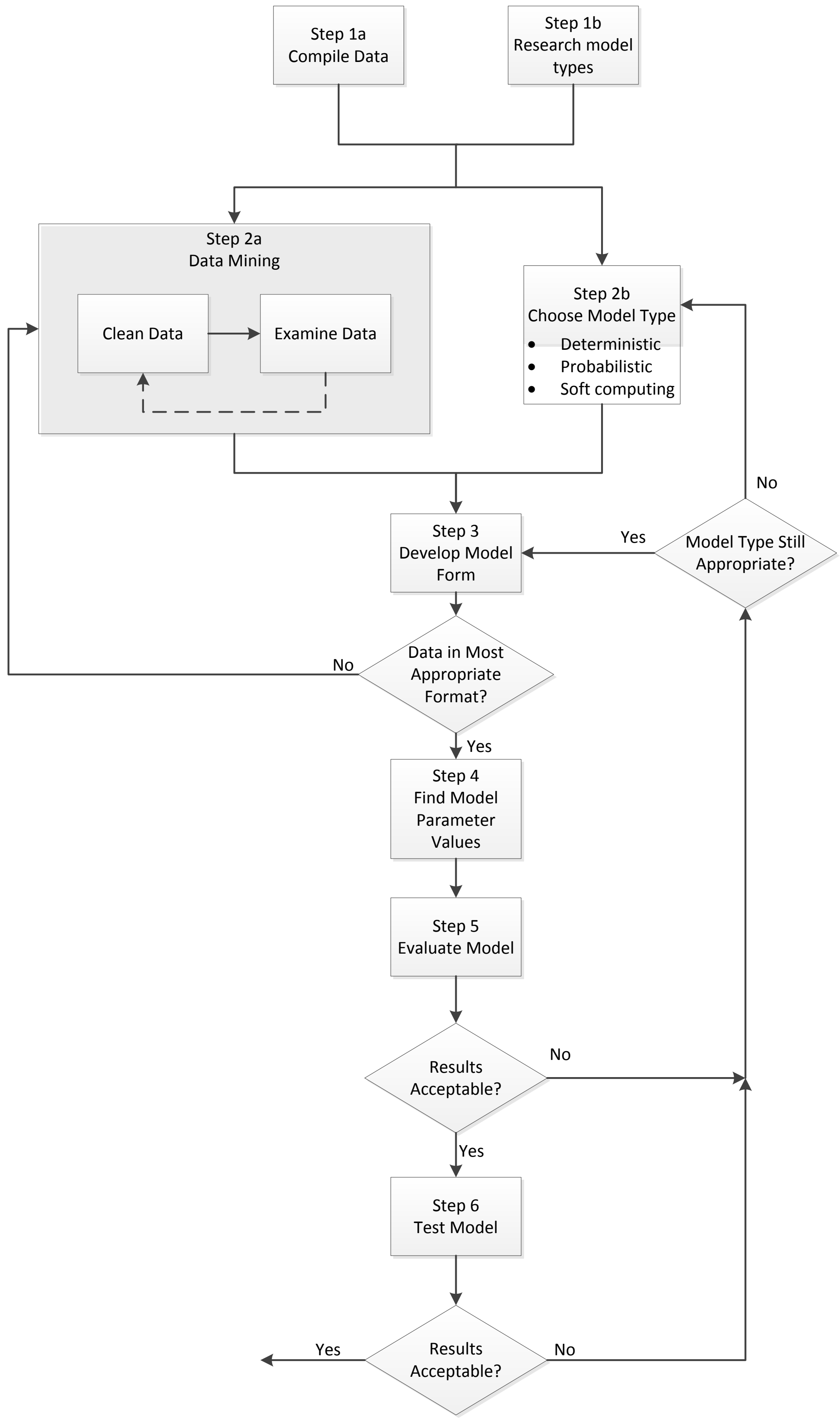


Figure 4 Proposed deterioration modelling framework

Several data sources may also be needed to develop an effective model.. For example, aggregated inspection data, digital soil maps, traffic data, pipe burst data, borehole logs and property age data have been combined and used to create a comprehensive dataset (Davies et al. 2001b). When historical records have not been transferred to digital files, this process can become very time-consuming and tedious. Possible data sources are found in Table 3.

Table 3 Possible data sources

Data Source	Possible Variable(s)
Inspection Data	condition, individual distresses
Asset Inventory	material, size, location, functional class, drainage information, slope
As-built Drawings	material, size, location, age, construction anomalies
Construction History	construction technique(s), construction standards, construction anomalies, contractor, inspector, age
Geographic Information System	location
Borehole Logs/ Soil Maps	soil types and properties
Traffic Data	present or forecasted AADT, ESALs
Failures Database	condition
Hydraulic Models	hydraulic loading, slope
Maintenance Records	applied treatment(s), age
Drainage/Sewerage Basin	quantity/quality of water/ waste
Climate/Weather Data	precipitation, temperature, freeze-thaw cycles

Some datasets include only “snapshot” data, or data that was taken at a single point in time, and only contains one data point per section (Baik et al. 2006; Ana et al. 2009; Tran et al. 2007). A problem that is theoretically applicable to all models where only a snapshot of data is available is survival bias (Le Gat 2008). In these models, the older an asset is, the slower its rate of deterioration. If the section deteriorated more quickly, it would have required rehabilitation and thus would not be classified as an older asset. This leads to an underestimation of the number of assets in the worst condition states, and a corresponding overestimation of the physical life-span of the asset. A complete historical record of the section would alleviate this problem in time-dependent models. In models independent of

time (such as the time-homogeneous Markov model), frequent inspection data has been found to decrease this bias (Scheidegger et al. 2011).

Even when longitudinal data are available, they are not necessarily accurate. When analysing sewer inspection data from the Netherlands, Dirksen (2008) found that many defects simply “disappeared” with time—defects identified at an initial inspection were not found at subsequent inspections.

Several recommendations have been made to work around these data challenges. Expert opinion has been suggested as a viable data source upon which a model can be developed until enough real data are obtained (Kleiner 2001). It has also been proposed that small communities could pool data if the deterioration factors amongst them were thought to be similar (Ana et al. 2010). Finally, condition data can be simulated to evaluate deterioration models (Scheidegger et al. 2011). Further information on compiling pavement data can be found in Chapter 4 of the Pavement Asset Design and Management Guide (TAC 2012).

3.2 Step 1b – Research Model Types

Deterioration models can be deterministic, probabilistic or based on soft computing. A brief overview of the models commonly used in infrastructure deterioration modelling can be found in Chapter 2. It is also useful to review technical papers and reports from agencies applying these models as well as research work done on the subject. The software requirements and technical expertise necessary to develop the model, and the types of data that are used for independent and dependent variables should be noted as models are reviewed.

3.3 Step 2a – Data Mining

This step involves cleaning and examining data that may be used in the deterioration model. When data is “cleaned”, incorrect and irrelevant data is removed from the dataset. This can include correcting typos, ensuring that data formats are consistent and removing records with incomplete data. The Oklahoma Department of Transportation (Wolters et al. 2006) developed an application that works with their pavement management system database to cleanse their data.

Variables can be measured on an interval, ordinal or nominal/categorical scale. Categorical variables are those that have no scale between variables and cannot be ordered. For example, environment, with categories urban, semi-urban and rural, is a categorical variable. To use categorical variables in a model, the categories must be converted to numbers. Dummy coding converts a categorical variable with n possible values into n-1 new variables with a value of 1 (present) or 0 (not present). For example, environment could be converted into two variables: semi-urban and rural. When semi-urban takes on a value of 1, it means the road section is semi-urban; when rural has a value of 1, the section is rural; and when both semi-urban and rural have 0 values, the section is urban. Categorical variables with only two values are called dichotomous.

Ordinal variables have a set order, but the degree of difference between values is not known. Condition is often measured on an ordinal scale. For example, the WRc condition assessment protocol for sewers assigns ratings from 1 (excellent) to 5 (collapsed). In this case, a sewer in condition grade 1 (excellent) is in better condition than a sewer in condition grade 2 (good), but the degree to which it is in better condition is unknown.

Continuous, or interval, data has a set order and the degree of difference between values is known. In infrastructure asset management, age is typically measured as a continuous variable. The greater the value, the older the asset, and the difference between values (1 year, 1 month, etc.) is the same for all values. Deterministic models often treat dependent variables measured on an ordinal scale as continuous.

Data binning and categorization are used to reduce the effect of minor measurement errors on the model, and to simplify data. Binning or categorizing variables is sometimes necessary to reduce the number of values when limited data is available.

Binning separates a continuous variable into several discrete groups. There are many methods of binning data. These methods include visually (based on breaks in the data), at equal intervals, based on a particular distribution, and optimized based on another variable.

Values within categorical variables can also be grouped together. Methods of grouping values include grouping based on expert knowledge (for example, that two pavement surface types

behave similarly), or grouping based on the difference between values with respect to another variable.

After the data has been through quality checks, it must then be examined. First, histograms or bar charts of the variables should be created to determine their frequency. Then, the correlation between variables should be found. Knowing the distributions and correlations between variables helps the modeller to choose an appropriate form for the data, and to more effectively evaluate the model once it has been created. As data is examined, inconsistencies and inaccuracies might be found signifying that further cleaning of the data is required.

Correlation refers to the relationship between two variables. A high correlation means that the two variables are closely related – as one variable changes, the other changes proportionally. If they are continuous variables, they form a line when plotted against each other. A very low correlation means that the two variables change randomly and are not associated. Most data fit somewhere between the two extremes

Measures of correlation differ depending on the level of measurement (interval, ordinal, etc.) of variables involved. Table 4 provides measures of correlation coeffients depending on the level of measurement of the variables involved.

Table 4 Measures of correlation

	interval	ordinal	nominal	dichotomous
interval	Pearson correlation coefficient (r^2)	Spearmann's ρ or Kendall's τ*	η (eta)***	point biserial
ordinal	Spearmann's ρ or Kendall's τ*	Spearmann's ρ or Kendall's τ	Contingency coefficient, Cramer's V**	rank biserial (somer's D)
nominal	η (eta)***	Contingency coefficient, Cramer's V**	Contingency coefficient, Cramer's V	Contingency coefficient, Cramer's V
dichotomous	point biserial	rank biserial (somer's D)	Contingency coefficient, Cramer's V	ϕ (phi)

* interval variable treated as ordinal
** ordinal variable treated as categorical
*** asymmetric measure

Further information on the calculation of these measures can be found in many statistics textbooks.

3.4 Step 2b – Choose Model Type

After considering each of the model types investigated earlier (deterministic, probabilistic and soft computing), the type of model must be selected. There are several factors that should be considered when choosing a model. These include:

- the nature of deterioration in the particular asset being modelled,
- the available data,
- the expected output from the model, and how the model will be used in its final form.

3.4.1 Nature of Deterioration in Model Selection

Different models have different underlying assumptions regarding the nature of deterioration. For deterministic models, one set of inputs always produces the same output. This implies that given the same set of independent variables, assets will always be in the same condition. Deterioration is generally modelled to occur gradually, with time.

Probabilistic models, such as logistic and probit regression, and Markov models, assume that the deterioration process is random, to some degree. Thus, a set of input data may not have the same outcome in every case. In time-homogeneous Markov models, the deterioration process is independent of time, whereas the semi-Markov model assumes that deterioration changes with time.

There are also assumptions about the nature of deterioration within models. One such assumption sometimes found in Markov models, is that a section can only deteriorate by one state in a given time period ΔT, if ΔT is small enough (Kleiner 2001; Wirahadikusumah et al. 2001). This assumption reduces the number of transitional probabilities to be calculated, but may not be valid for all types of deterioration.

For example, in a sewer deterioration model, a single damage event (e.g. a very heavily loaded truck) could cause a pipe to structurally deteriorate several states in a given time

period (Micevski et al. 2002). These types of events would seem to be best described using a probabilistic model.

3.4.2 Data and Model Selection

If limited data are available, certain types of models are difficult, if not impossible, to create. Of the model types discussed, soft computing methods generally require the largest datasets (1000s of data). For these methods; the larger the dataset is, the longer the required time for training the model.

"Snapshot" data, for instance, can also potentially preclude several model types. If this is all the data available, it can be particularly problematic when some models, such as Markov models, require the condition of the segment at multiple time points (longitudinal data).

Also, due to the fact that mechanistic-empirical models usually require some form of measurement (deflection, strain, etc.) as an input to the model, certain data (e.g. rut depth, extent of cracking) must be collected and stored. If this detailed condition information is not available, the model cannot be used.

3.4.3 Expected Output

It is important to consider how a deterioration model will be used in its final form. For planning and budgeting purposes, a long-term forecast of the network is preferred, but the deterioration model for an individual asset segment is unnecessary. For inspection scheduling and insight into the deterioration processes, a deterioration model for an individual section is preferred (MTO 1991). Deterministic models result in an individual segment's deterioration curve (Chughtai et al. 2008). Markov models can output either individual section condition (Baik et al. 2006; Le Gat 2008) or the condition of the network (Wirahadikusumah et al. 2001; Micevski et al. 2002) and neural networks typically output the individual segment condition (Najafi et al. 2005; Tran et al. 2007).

The integration of the deterioration model with the overall asset management system can also affect the choice of model. For example, if an asset management system makes decisions based on risk, a probabilistic deterioration model may be better suited than a deterministic one.

It may be important that the deterioration model is easy to use and understand. In this case, transparency in the model development process and in how the model is used after it has been developed is essential. Soft computing techniques are generally the most complex to set up and to use after their development process is complete. It is often difficult to see and understand how the inputs produce outputs in these model types.

3.4.4 Other Considerations

Other factors that should be considered by the modeller include any underlying assumptions related to the model type. For example, for deterministic models, it is assumed that the variables are distributed multivariate normal in the population (i.e. that each variable is itself normally distributed and is also normally distributed for any possible combination of other variables). For survival analysis, the proportional hazards assumption states that independent variables have the same effect on condition regardless of an asset's initial condition state. This is similar to the proportional odds assumption necessary for ordinal logistic regression. While this and other assumptions related to a particular model types must be considered, they do not necessarily preclude a model type from being selected. If an agency, for example, is simply attempting to find the best model to fit their data, and will not be extrapolating or applying the model to other datasets, how well a model predicts condition may be more important than ensuring statistical assumptions are met.

The level of complexity and how this relates to the time available for development should also be considered. Deterministic models are the simplest and least time-consuming model discussed. These models may not require more than a spreadsheet for their development. Soft computing methods are the most complex and time-consuming to create.

3.4.5 Summary of Decision Process

Choosing a deterioration model is not a straight-forward process. The choice of model depends on many factors and the importance the decision-maker attributes to these factors. Figure 5 illustrates the potential decision process. It should be noted that this figure does not include all of the model types that are relevant to infrastructure deterioration modelling, and a model that is not selected using the process may still produce reasonable results. Many of the

decisions to be made using this process are discussed earlier in Chapters 2 and 3. The questions found in Figure 5 are presented in no particular order.

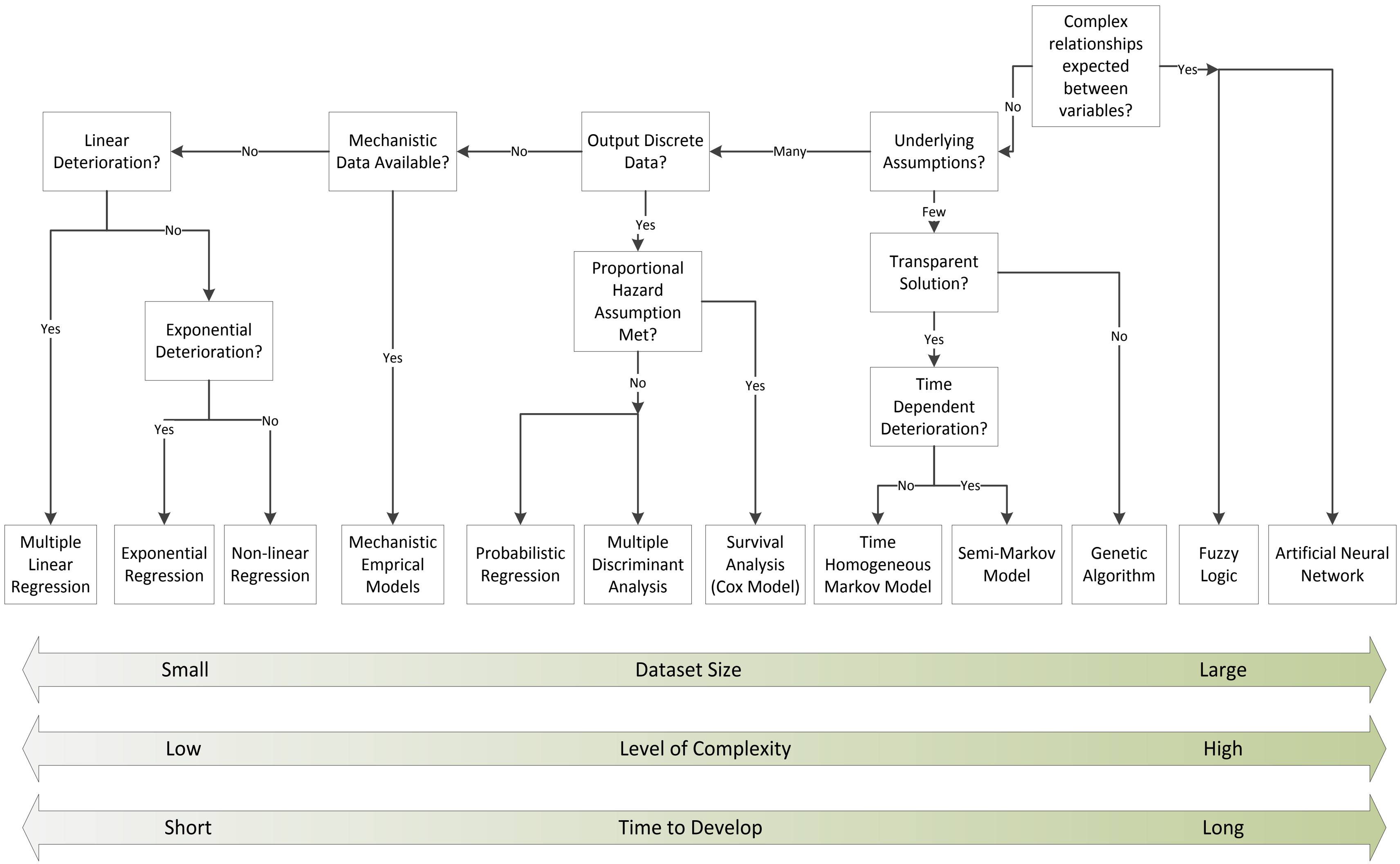


Figure 5 Summary of decision process

3.5 Step 3 – Develop Model Form

Creating the best deterioration model is usually an iterative process, with the modeller changing various aspects of the model form to create the best model using the available data. The aspects of the model form that can be changed vary by model type and by the software used to develop the model.

In the case of deterministic model types, some factors, among others, that affect the model form include the base equation, the variables used in the model, and how those variables are grouped. The base equation is the equation from which parameter values are optimized. In linear models, this does not typically need to be specified in the modelling software because the equation is known to be linear. In non-linear models, however, the exact form of the base equation does need to be specified and can be changed in successive iterations.

Of all the variables initially thought to be used in the model, only some will be found statistically significant. For some model types (stepwise regression comes to mind), variables will be automatically added to, or removed from, the model based on their significance. Other model types may require that the variables used in the model be entered, and manually changed, for each iteration.

Sometimes, in both probabilistic and deterministic models, assets are grouped into “families” of similar types. Separate analyses are then performed on each family group. Changing the members of these families can be thought of as changing the form of the model.

Soft computing methods require other factors to be pre-determined by the modeller. Neural networks, for example, require the number of hidden layers, and the type of training algorithm to be specified by the user.

It may be found that with certain models forms, data must be reformatted or regrouped to be used in the model’s development. Also, some models may be developed with a fraction of the total dataset. The remaining data will be used to evaluate the model.

3.6 Step 4 – Model Development

Developing the model entails finding the parameter values. Usually this is completed using an optimization algorithm, but for very simple models (a linear regression model using the method of least squares, for example) these values may be optimized manually using a spreadsheet program.

3.7 Step 5 – Model Evaluation

Once the parameter values have been found, and the model has been created, the model must be evaluated. The method of evaluation will depend on the model type. If the model evaluation is not deemed acceptable, the model type should be reconsidered. If the model type is still thought to be appropriate, the model form should be changed and the model re-developed. If the results of the evaluation lead to the conclusion that the model type is not appropriate for the available data, the type of model should be reconsidered.

Many measures can be used to evaluate statistical models. One of the first measures that should be considered when evaluating a model is the parameter estimates. The parameter values should be reasonable and significant.

A significant parameter value means that the associated independent variable explains a significant variation in the dependent variable given the presence of the other independent variables in the model. Significance is measured as a p-value on a scale of 0 to 1, and/or is shown as a confidence interval. Generally, a low p-value (less than 0.05 or 0.01), and a confidence interval that is relatively small and does not bridge 0, means that the parameter is significant.

Whether a parameter value is reasonable or not is based on prior knowledge. For example, it would make sense that a parameter value associated with age should be negative (since condition decreases as age increases), and relatively small (given the relative scale of condition to age).

Another method that can be used to evaluate any predictive model (deterministic, probabilistic or soft computing) is a plot of the residuals over the dependent variable. The residual for each data point is calculated by subtracting the value of the dependent variable

from the predicted value. If the residuals have similar values across the dependent variable, the model is said to be homoscedastic. If residuals are not homoscedastic, the model is better at predicting over certain intervals of the dependent variable, and any conclusions drawn from the R^2 statistic (described below) should be considered in light of this.

An important statistic when evaluating a model is the coefficient of determination, R^2. R^2 is a measure of how much of the variance in structural adequacy is described by the model, or, how well the model fits the actual data. Depending on the model type, R^2 can be calculated in different ways. This means that the R^2 values of different model types cannot be directly compared.

In most cases, R^2 ranges from 0 to 1. Zero means that none of the variation in the data is explained by the model – a very poor fit, and 1 means that all the variation in the data is explained by the model – a perfect fit. For deterministic models, R^2 is calculated as 1 – (residual sum of squares)/(corrected sum of squares).

Finally, any assumptions that were made when creating the model should be evaluated. It should be noted, however, that a violation of the assumptions may not be cause to discard the model if it still performs relatively well.

It is not possible to explore the evaluation procedures for all the model types presented in Section 2.3. However, since ordinal regression models are presented in the case studies presented in this paper, the evaluation procedures for this model type are presented.

3.7.1 Evaluative Measure Specific to Ordinal Logistic Regression

There are three pseudo-R^2 values that are commonly used to measure the strength of association between the dependent and independent variables in ordinal logistic regression. These values cannot be interpreted in the same way as the R^2, which is used in the linear and exponential models.

Cox and Snell R^2:

$$R^2{}_{CS} = 1 - \left(\frac{L(B^{(0)})}{L(\widehat{B})}\right)^{\frac{2}{n}} \qquad [7]$$

Nagelkerke's R^2:

$$R^2{}_N = \frac{R^2{}_{CS}}{1-L(B^{(0)})^{2/n}} \quad [8]$$

McFadden's R^2:

$$R^2{}_M = 1 - \left(\frac{L(\widehat{B})}{L(B^{(0)})}\right) \quad [9]$$

Where $L(\hat{B})$ is the log-likelihood function for the model with the estimated parameters, $L(B^{(0)})$ is the log-likelihood with the parameter values set to 0, and n is the number of data points or observations.

Another method commonly used to determine if an ordinal regression model fits the data is the difference between the log-likelihood of the model with parameter values (except the intercept values) set to 0 and parameter values as set by the model. If the difference is significant, then the model with predictors is better than the model without them.

When creating an ordinal logistic model, proportional odds are assumed. This means that the relationship between independent variables and the log-odds is the same for all values of the dependent variable. The test of parallel lines, which compares the -2 log likelihood in the case that the relationship is the same to the case where the relationship is not the same, tests this assumption.

The fit of an ordinal logistic regression model also can be evaluated using Pearson and Deviance goodness-of-fit measures. These measures relate to the observed and expected frequencies for each of the category combinations. The Pearson goodness-of-fit statistic and deviance measure can be found below.

$$\mathrm{X}^2 = \sum\sum \frac{(O_{ij}-E_{ij})^2}{E_{ij}} \quad [10]$$

$$D = 2\sum\sum O_{ij}\ln\left(\frac{O_{ij}}{E_{ij}}\right) \quad [11]$$

Where O_{ij} is the observed frequency, the actual number of road sections in a particular structural adequacy category and set of independent variables; and E_{ij} is the expected frequency, the predicted number of road sections in a particular structural adequacy category and set of independent variables.

Where there are many combinations of variables with low or 0 expected cell counts, these goodness-of-fit statistics may not give accurate results.

3.8 Step 6 – Test Model

A model can be theoretically very good according to its evaluative statistics and methods, and still not be the best, or even a good, overall choice (Ahammed et al. 2008). Testing the model involves checking that it is suitable for its intended purpose. Testing can involve comparing the model to expected results or to other models developed with the same dataset. It involves ensuring that the model works within the overall asset management system and that it performs as expected in critical ranges.

If the model testing is not deemed acceptable, the model type should be reconsidered. If the model type is still thought to be appropriate, the model form should be changed and the model re-developed. If the results of the evaluation lead to the conclusion that the model type is not appropriate for the available data, the type of model should be reconsidered.

3.9 Summary

In this chapter, the proposed framework for creating a deterioration model is presented. The steps to creating a model are:

Step 1a – Compile Data: Compile data (possibly from multiple sources) to be used as variables in the model.

Step 1b – Research Model Types: Review possible models noting the software requirements and technical expertise necessary to develop the model, and the types of data that are used for independent and dependent variables.

Step 2a – Data Mining: Clean and examine data to be used to create the model.

Step 2b – Choose Model Type: Choose a model type considering the nature of deterioration in the particular asset being modelled, the available data and the expected output, and how it will be used in its final form. An example of a potential decision process is presented.

Step 3 – Develop Model Form: Change various aspects of the model (e.g. the base equation, set of independent variables) to eventually find the form that best suits the data.

Step 4 – Develop Model: Find parameter values for the model.

Step 5 – Evaluate Model: Evaluate how well the model model fits the data. The method of evaluation will depend on the model type.

Step 6 – Test Model: Check that the model is suitable for its intended purpose.

4 Case Study: Pavement Deterioration Model

This chapter presents an application of the framework outlined in Chapter 3, applied to a dataset from the City of Oshawa. The City is creating a new, risk-based asset management plan. The pavement deterioration models used in their previous plan were based on expert opinion rather than actual inspection results. The City of Oshawa has a full set of inspection data from 2009, when the entire road network was inspected. The aim of this case study is to present three deterioration models that have been developed using the City of Oshawa's data. A discussion as to which model would best suit the City's needs is provided.

4.1 Step 1a – Compile Data

The data for the proposed models were provided by the City of Oshawa. The final data set was assembled using Microsoft Access. It contains approximately 1700 records; one record for every unique segment of road. A unique segment of road has:

- One road segment identifier (RDSEC),
- Uniform construction, maintenance and rehabilitation activities; and
- A unique inspection record per inspection cycle.

Only records that contained information in every field were used in the analysis. Those records that showed no maintenance or rehabilitation of a road segment in the last 35 years were excluded from the analysis as it was assumed some treatment or activity had not been recorded, making that record unreliable. From these data, several independent variables, or factors suspected to influence deterioration, were extracted. Structural adequacy scores, collected via inspection in 2009 for the entire road network, were used as the dependent variable for the models. Figure 6 outlines the possible variables for the model.

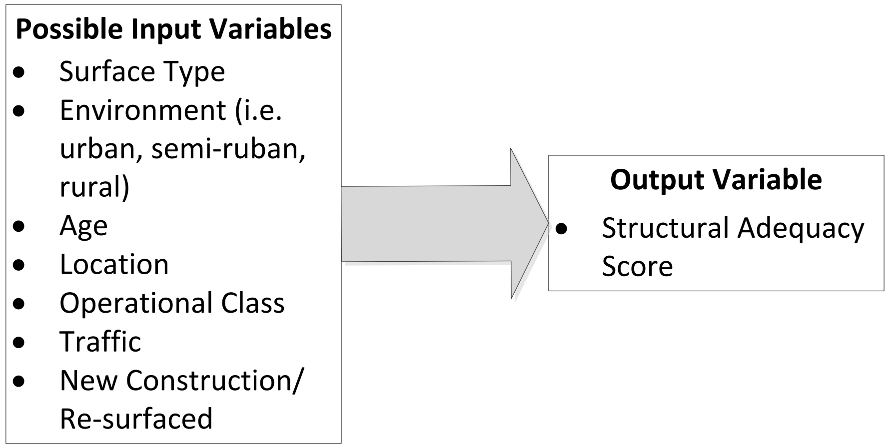


Figure 6 Pavement deterioration model variables

4.2 Step 1b – Research Model Types

An overview of the model types used to model pavement deterioration is found in Chapter 2.

4.3 Step 2a – Data Mining

The following section provides an overview of the variables used in the deterioration models.

4.3.1 Dependent Variable – Structural Adequacy Score

Structural adequacy is a score that describes the structural condition of a pavement section. In the City of Oshawa, road sections are visually inspected and assigned a structural adequacy score from 0 (structurally inadequate) to 20 (structurally adequate). The inspector evaluates the severity and extent of pavement distresses (e.g. cracking, ravelling, rutting, etc.) on a particular section, and assigns a grade based on when they expect work will be needed on the section. Table 5, taken from the Ontario Road Needs Manual (1991), provides the range of years when work is expected to be necessary corresponding to the structural adequacy score.

Based on the 2009 inspection of the network, Figure 7 shows the distribution of structural adequacy scores. 86% of the road segments are in good/adequate condition according to Table 5, with scores between 15 and 20, while 7% have a score between 12 and 14 and the remaining 7% have scores between 0 and 11.

Table 5 Structural adequacy scores

Structural Adequacy Score	Years in which work will be necessary
15-20	Currently in adequate condition
12-14	6-10
8-11	1-5
0-7	Currently needs work

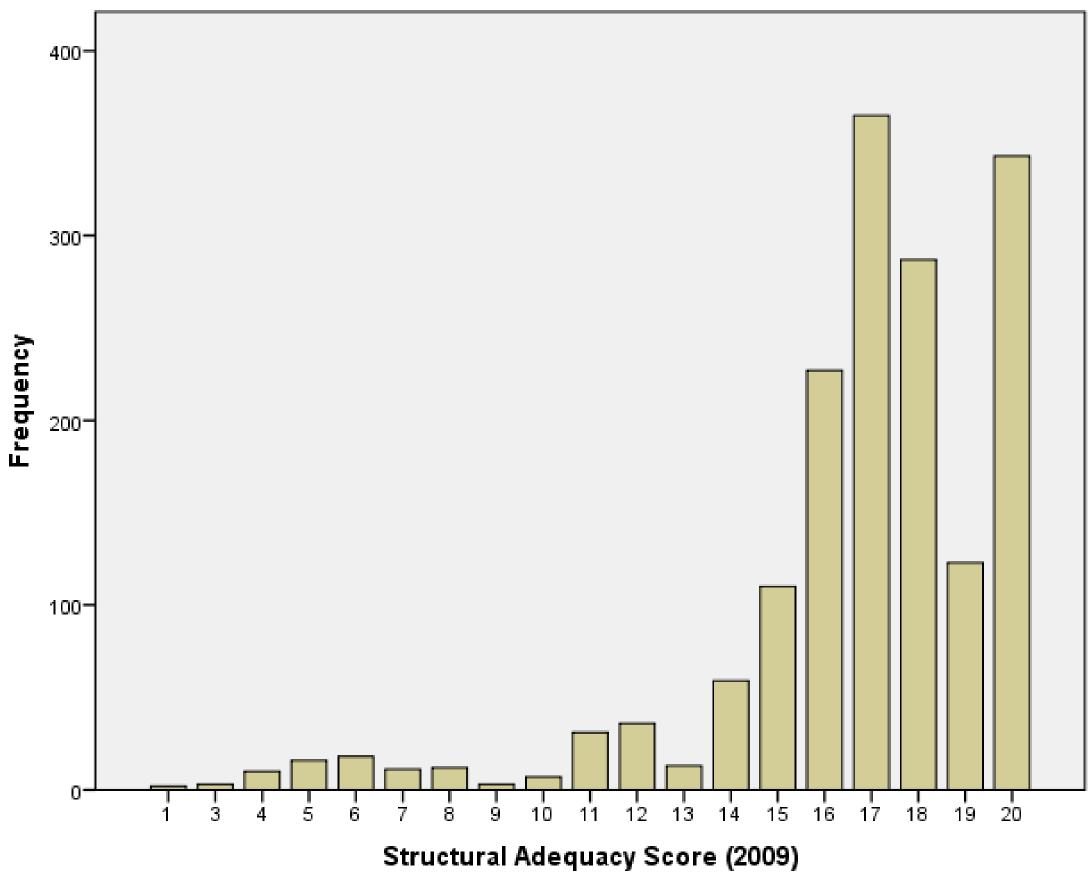


Figure 7 Structural adequacy score histogram

To create the simplest model possible while still including all the information necessary for decision-making, structural adequacy, originally measured on a scale of 0 to 20, was transformed into an ordinal scale of 1 to 7 for use in the ordinal logistic regression model.

The ranges for these groups were created using the “trigger” values for maintenance treatments defined by the City of Oshawa so that the ordinal regression model could be used in conjunction with the City’s asset management software. The ranges for the binned structural adequacy scores are shown Table 6.

Table 6 Binned structural adequacy scores (ordinal logistic model only)

Original Structural Adequacy Score	New Structural Adequacy Category
0-7	1
8-10	2
11	3
12	4
13	5
14	6
15-20	7

It should be emphasized that the new structural adequacy categories (as well as the original scores) are measured on an ordinal scale. Thus, a pavement in category 6 is not in 2 times better a condition than one in category 3.

4.3.2 Independent Variables

4.3.2.1 Surface Material

Surface material refers to the material used in the surface course of the pavement. Table 7 provides the surface types that are found in the City of Oshawa’s road network. (Gravel roads were not included in this analysis.)

Table 7 City of Oshawa surface types

Surface Material	Frequency	Percent
Asphalt over Concrete (A/C)	2	.1
High Class Bituminous (HCB)	1593	95.0
Intermediate Class Bituminous (ICB)	5	.3
Low Class Bituminous (LCB)	75	4.5
Prime (PRI)	1	.1
Total	**1676**	**100.0**

The network is primarily (95%) composed of a high class bituminous surface course.

To decrease the number of variables included in the analysis, several of the categories (e.g. A/C, HCB, etc.) were binned together. The groups were found by ordering the categories by their mean structural adequacy score. Then, the Mann-Whitney U test was used to determine if adjacent categories had significantly different structural adequacy scores. Those categories that were not significantly different from one another were binned together. Using this method, data was separated into two groups. The first includes pavements with A/C, HCB and ICB surface course, and the second group includes LCB and PRI pavements.

4.3.2.2 Environment

Environment refers to the surrounding land use of the road segment: rural, urban or semi-urban. This variable was extracted directly from the City of Oshawa’s data. Table 8 provides the environment types found in the data set:

Table 8 Environment types

Environment	Frequency	Percent
Rural	78	4.7
Semi-Urban	32	1.9
Urban	1566	93.4
Total	**1676**	**100.0**

The majority of the sections are in an urban area. Because the environment variable only has 3 categories, it was not thought necessary to group it any further.

4.3.2.3 Age of Infrastructure

The value used for the age of the pavement was given careful consideration. Three separate “ages” were proposed: the age of the surface course, the age of the base, and the age based on construction. Considering the fact that the structural adequacy score for a segment is reset whenever any type of maintenance activity (except crack sealing) has occurred, it was thought that the age of the surface course would be most appropriate as an independent variable. However, base age has been considered in the variable “New Construction/ Rehabilitated”, as described below. From Figure 8 and Figure 9 it can also be seen that there is a much stronger downward trend in the average structural adequacy score when the age of the surface course is considered rather than the age of the base.

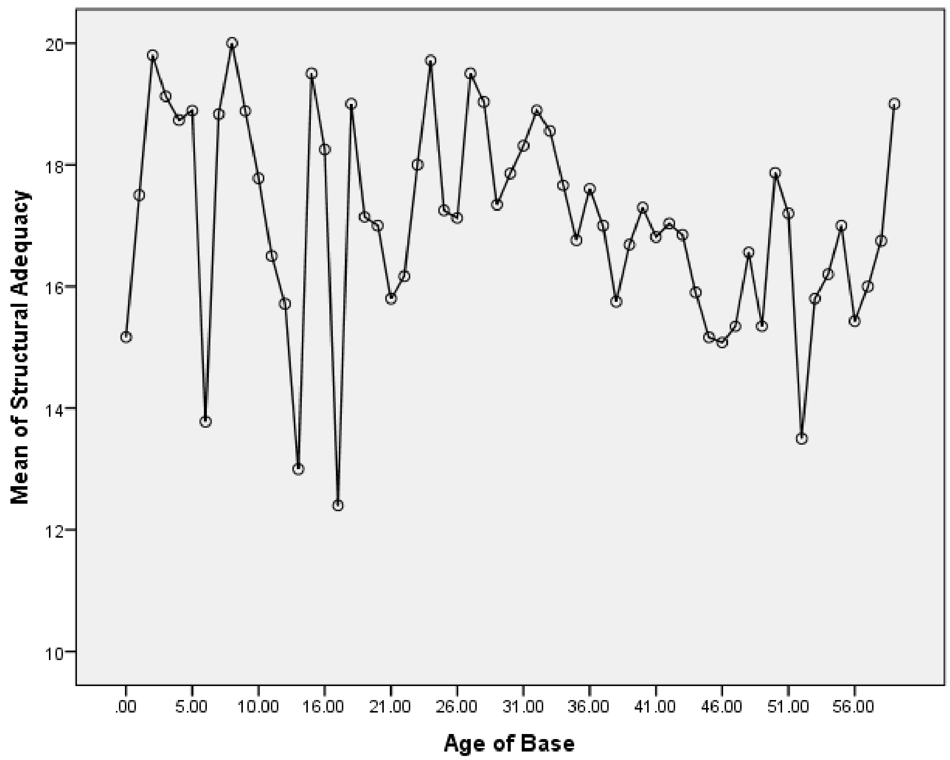


Figure 8 Mean structural adequacy score by base age

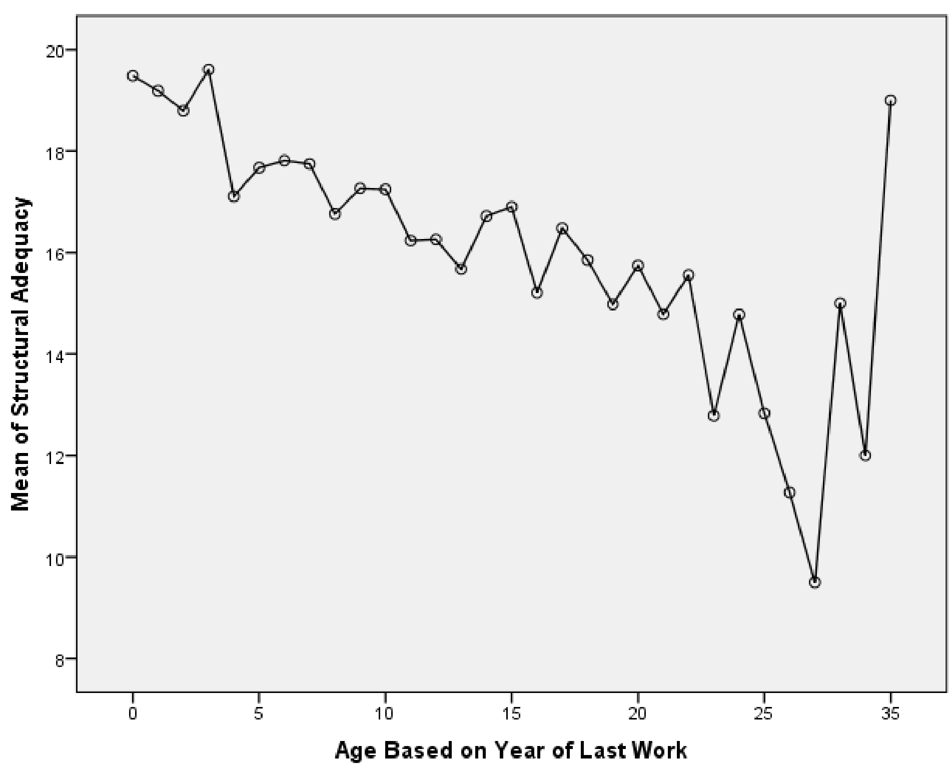


Figure 9 Mean structural adequacy score by age (based on year of last work)

Only road segments with ages 35 or less were used in this analysis, as it seems likely that maintenance or rehabilitation of segments over 35 years had occurred but was not recorded.

Figure 10 shows the distribution of sections based on age.

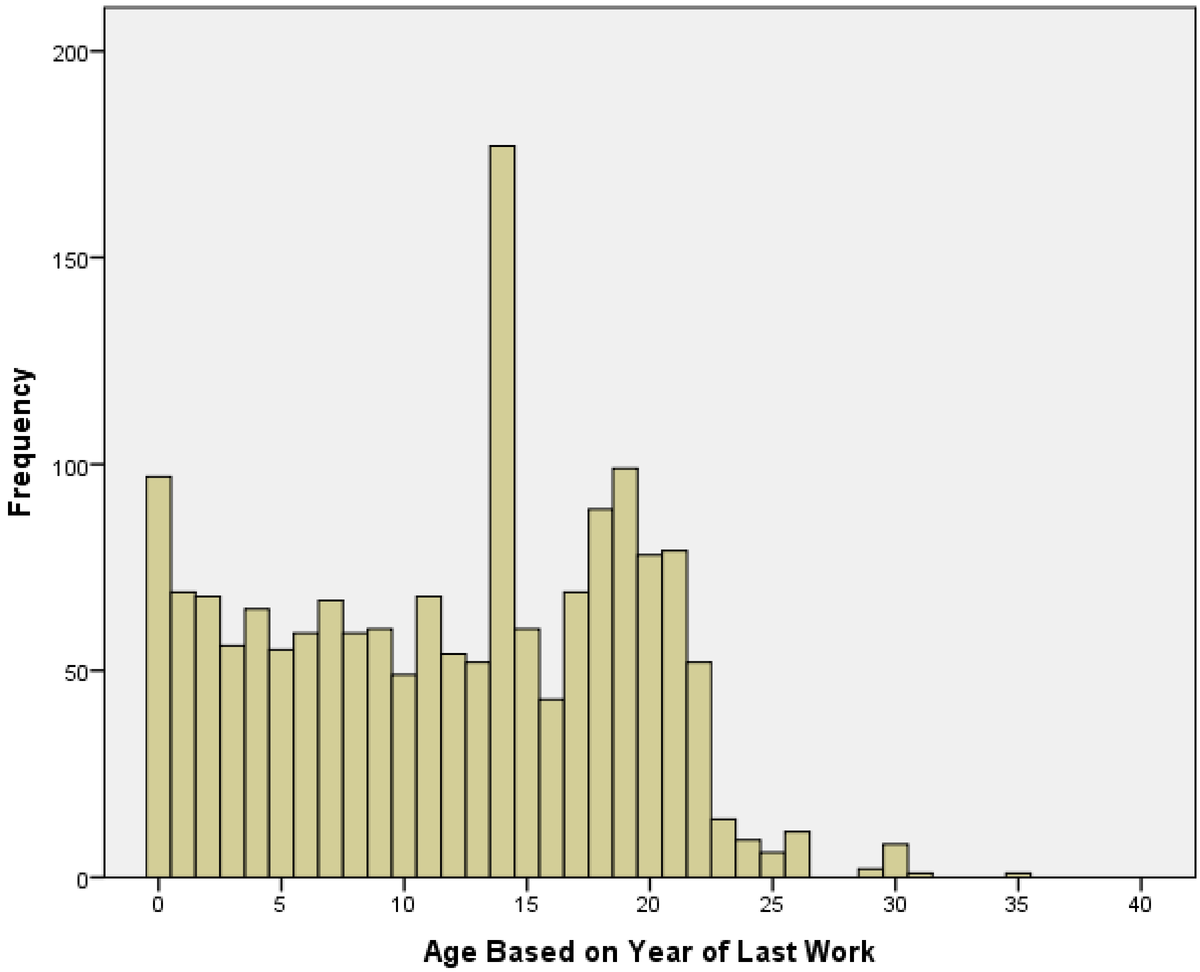


Figure 10 Histogram of road segments by age (based on year of last work)

Many rehabilitation and maintenance treatments took place in 1995 (or 14 years prior to 2009) due to an increase in provincial funding. Other than age 14, it can be seen that the distribution is generally fairly uniform until around age 22, after which there are fewer data points. This is because many sections had deteriorated to such a point that rehabilitation was necessary – consequently the age of the section was reset.

Because age as a continuous variable was found to be insignificant in the first attempt at an ordinal regression model, the variable was binned to produce significant results. Road sections were assigned an age “group”, which was used instead of age in the analysis. Table 9 shows the age groups and their corresponding age ranges.

Table 9 Binned age groups

Age Range	Age Group
0-4	0
5-9	1
10-14	2
15-19	3
20-24	4
25-29	5
30-34	6

4.3.2.4 Location/Area

Each road segment is located in a particular area 1 through 9 as defined by the City of Oshawa. A sketch of these areas can be found in Figure 11. These areas were then binned together based on the differences in their structural adequacy scores. The areas were first ranked, and then analysed in pairs to determine if the structural adequacy scores were significantly different using the Mann-Whitney U test. It was found that there was no significant difference in the structural adequacy scores between areas (at alpha = 1%) except between areas 1 through 7 and areas 8 through 9. Areas 8 and 9 are in the northern half of the City of Oshawa. Thus, the location/area variable can take one of two values: area 8 or 9 (~1200 records), or areas 1 through 7 (~500 records).

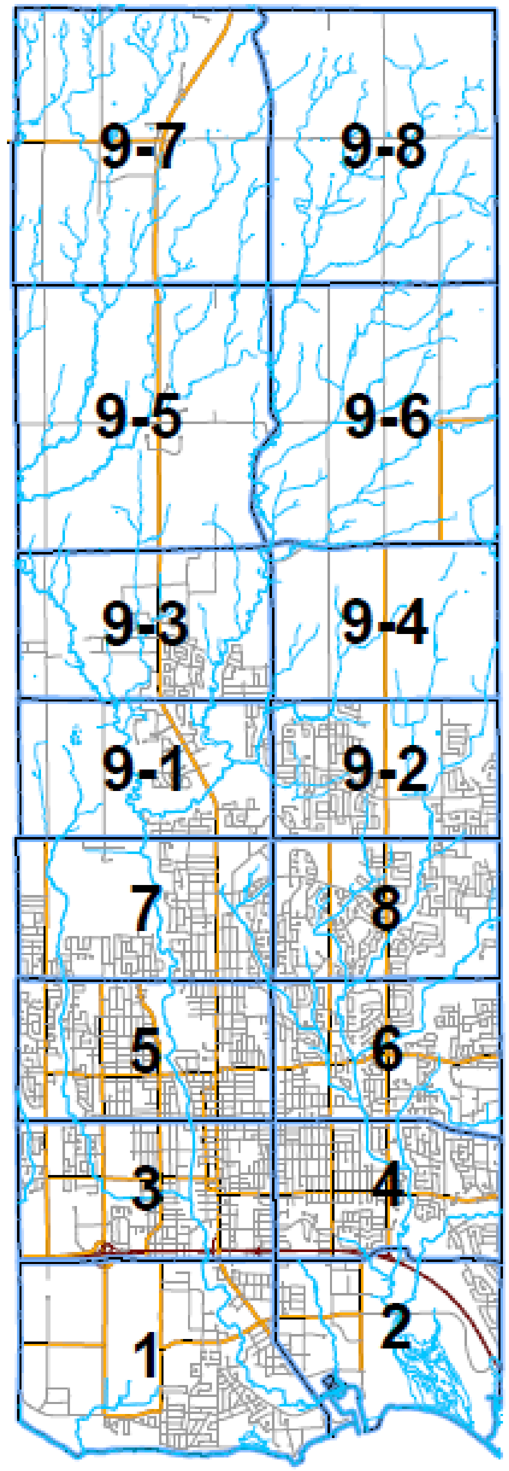


Figure 11 Sketch of City of Oshawa locations

4.3.2.5 Class

Class is the operational class of the road: local, collectors, or arterial (Hwy Class A, Hwy Class B or Hwy Class C). The details on this classification for the City of Oshawa can be found in Table 10. Similarly to the analysis performed on the location/area and surface type variables above, the class variable was binned according to the difference in structural adequacy between groups. It was found that there was a significant difference between all groups except local and collector roads. Thus, four groups were created: local and collectors (as one group), Highway Class A, Highway Class B and Highway Class C.

Figure 12 shows that most of the sections are local or collector roads and very few are in Hwy Class A.

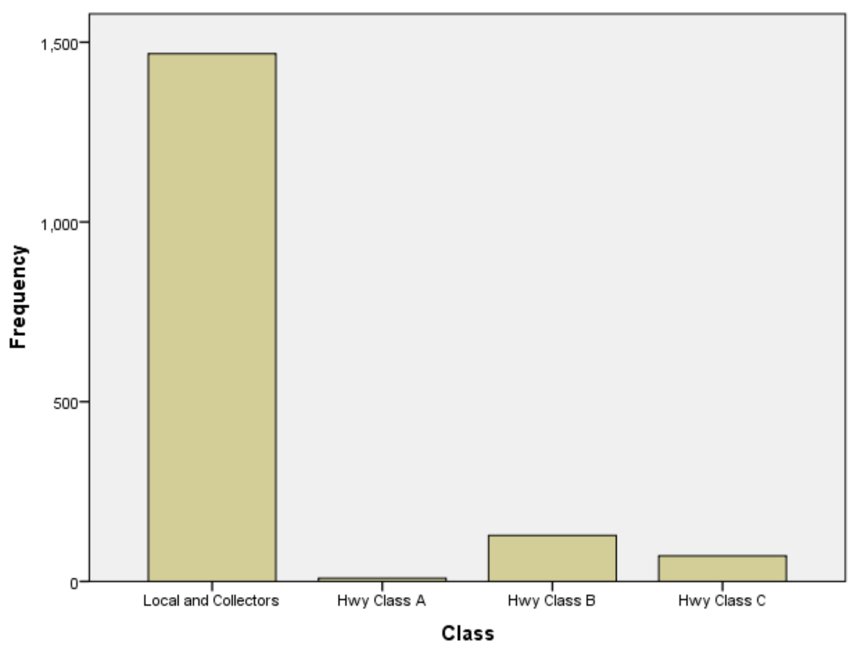


Figure 12 Class distribution

Table 10 Classification of City of Oshawa roads

ROAD TYPE	GENERAL FUNCTION	TYPICAL RIGHT-OF-WAY WIDTH	INTERSECTION AND ACCESS
Type "A" Arterial	Large volumes of traffic, including large volumes of truck traffic	36m to 50m (118 to 164 ft.)	Generally only intersect with freeways and other arterial roads to provide highest level of service and may accommodate high occupancy vehicle or bus lanes where required. Direct access to adjacent property to be controlled or not permitted. Generally private accesses shall be located a minimum of 200m (655 ft.) apart in urban areas.
Type "B" Arterial	Moderate volumes of traffic, including moderate volumes of truck traffic	30m to 36m (98 to 118 ft.)	Generally intersect with other arterial and collector roads to provide a moderate level of service and may accommodate high occupancy vehicle and bus lanes where required. Direct access to adjacent property to be controlled or not permitted. Generally private accesses shall be located a minimum of 80m (260 ft.) apart in urban areas.
Type "C" Arterial	Lower volumes of traffic including lower levels of truck traffic.	26m to 30m (85 to 100 ft.)	Generally intersect with Type "B" and Type "C" arterial and collector roads. Direct access to adjacent property will be permitted subject to acceptable crossing, stopping and sight line distances.
Collector	Moderate volumes of short distance traffic and light or moderate volumes of truck traffic moving between points of origin and arterial roads including local truck traffic.	(a) Urban - 20m to 26m (66 to 85 ft.) (b) Rural - 30m (98 ft.)	Generally intersect with Type "B" and Type "C" arterial, collector and local roads. Direct access to adjacent property will be permitted subject to acceptable crossing and stopping sight distances.
Local	Light volumes of traffic moving between points of origin and the collector road system.	(a) Urban - 20m (66 ft.) (b) Rural - 30m (98 ft.)	Generally intersect with collector, Type "C" arterial and local roads. Direct access to adjacent property to be permitted. Intersection of local roads with arterial Type "A" and Type "B" arterial roads is to be discouraged.

4.3.2.6 Traffic

Equivalent Single Axle Loads (ESALs) are generally accepted as a way to represent the damage to a pavement from its traffic loading. ESALs are typically calculated as a percentage of average annual daily traffic (AADT). A fraction of AADT is found based on the percentage of heavy vehicles, the type of heavy vehicles, and possibly other factors (e.g. % of trucks in design lane, traffic days per year, etc.). For example, the SHRP program uses the following equation to calculate ESALs for a 1 lane road (in one direction) (Haas 1997).

$$ESALs = 182.5 * AADT * TP * TF \quad [12]$$

Where AADT is the average annual daily traffic, TP is the percentage of heavy vehicles and combinations, and TF is the truck factor (which varies by region, pavement type, and type of truck).

Since the only data available from the City of Oshawa relating to ESALs were AADT, truck % and whether buses were present (buses were not included in the overall truck %), it was not possible to calculate ESALs from an existing equation (such as the SHRP method shown above). Therefore, traffic was calculated as an equivalent to ESALs. In this case, the following equation is used to as a simple means to compare ESALs between sections:

$$ESAL \propto AADT * (HVP + (0.01\ if\ buses\ present; 0\ if\ buses\ not\ present\ )) \quad [13]$$

Where AADT is the average annual daily traffic in 2009 assuming a linear growth factor, and HVP is the heavy vehicle percentage. Because no information was provided as to the quantity of buses present, 1% of the overall AADT was assumed. Figure 13 shows a histogram of the equivalent ESALs.

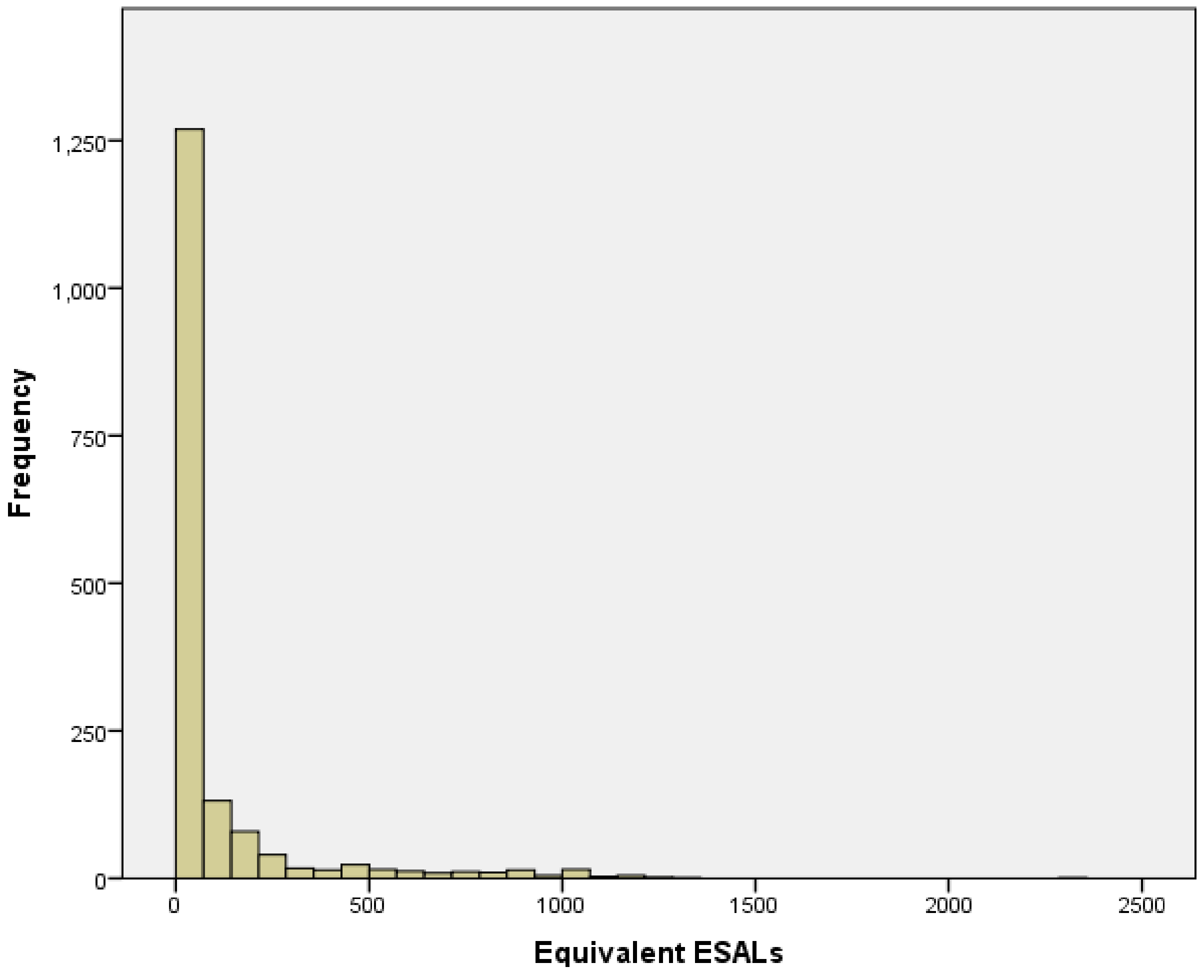


Figure 13 Histogram of equivalent ESALs

Most of the sections were found to have an equivalent ESAL of 0. This is because many sections had a heavy vehicle percentage of 0, and were not part of a bus route. While it is true that these sections likely have some damage from traffic (even light vehicles damage pavement – although to a lesser extent than heavy vehicles), for the purposes of this study the equivalent measure is sufficient because only a comparison between road sections is required – not an actual ESAL value.

The equivalent ESALs were then binned into 3 groups using the optimal binning technique (with structural adequacy as a guide variable) in SPSS. The upper and lower bounds for the bins are shown in Table 11.

Table 11 Binned equivalent ESALs

Bin	Lower Bound	Upper Bound	Number of Sections
1	0	18.9	1114
2	19	141.9	280
3	142	+	281

4.3.2.7 New Construction/Re-surface

The structural adequacy of a road section varies depending on whether it is newly constructed or has only been resurfaced. The variable is binary with 1 representing a pavement that has been re-surfaced and 0 representing a pavement that is a new construction. It was assumed that if the surface course was constructed within 2 years of the base, the pavement is a “new” construction; otherwise the section is classified as re-surfaced. 159 sections have been classified as new, and the remaining 1517 are classified as re-surfaced.

4.3.3 Correlations

Correlation is used to describe how much one variable’s value depends on another variable. A high correlation means that the two variables are associated – as one variable changes, the other changes as well. A low correlation means that the two variables are not associated – as one changes, the other does not. A summary of the correlation statistics for the variables discussed above can be found in Table 12.

Most of the variables have little or no correlation to one another. However, there are several variables that are moderately to strongly correlated. Age and structural adequacy have a moderate correlation. This appears to make sense logically, pavement deterioration generally does occur with age. Surface type and structural adequacy are also strongly correlated, so it is not surprising that surface type is significant in the models. Surface type and environment have a strong correlation. This makes sense, as rural roads are likely to have a less expensive surface since their traffic levels are lower. Class and equivalent ESALs are moderately correlated, since class is partly based on traffic levels.

Table 12 Correlation statistics

	Age	Age_bin	STAD	STAD_bin	ESAL_groups	surface_type	area	environment	class
Age_bin	N/A								
STAD	Kendall's tau-b: -0.499, Spearman's rho: -0.624	Kendall's: -0.514, Spearman's: -0.613							
STAD_bin	Kendall's tau-b: -0.196, Spearman's rho: -0.239	Kendall's: -0.195, Spearman's: -0.223	N/A						
ESAL_groups	eta: 0.309	eta: 0.210	Spearman's: -0.145, Cramer's v: 0.296	Spearman's: -0.216, Cramer's v: 0.214					
surface_type	Point bi-serial: -0.064	Point bi-serial: -0.067	rank bi-serial: -0.811	rank bi-serial: -0.718	Cramer's v: 0.181, Contingency coeff: 0.178				
area	Point bi-serial: 0.05	Point bi-serial: 0.048	rank bi-serial: 0.212	rank bi-serial: 0.112	Cramer's v: 0.134, Contingency coeff: 0.133	phi: 0.306			
environment	eta: 0.280	eta: 0.140	Spearman's: -0.219, Cramer's v: 0.471	Spearman's: -0.380, Cramer's v: 0.435	Cramer's v: 0.157, Contingency coeff: 0.216	Cramer's v: 0.852, Contingency coeff: 0.649	Cramer's v: 0.334, Contingency coeff: 0.317		
class	eta: 0.282	eta: 0.149	Spearman's: -0.227, Cramer's v: 0.306	Spearman's: -0.291, Cramer's v: 0.238	Cramer's v: 0.569, Contingency coeff: 0.553	Cramer's v: 0.370, Contingency coeff: 0.347	Cramer's v: 0.164, Contingency coeff: 0.162	Cramer's v: 0.321, Contingency coeff: 0.414	
new/resurf	Point bi-serial: 0.113	Point bi-serial: 0.115	rank bi-serial: -0.248	rank bi-serial: -0.023	Cramer's v: 0.036, Contingency coeff: 0.036	phi: -0.125	phi: -0.33	Cramer's v: 0.235, Contingency coeff: 0.229	Cramer's v: 0.125, Contingency coeff: 0.124

- Very Strong Correlation (> 0.7)
- Moderately Strong Correlation (0.5 – 0.7)
- Weak Correlation (0.2 – 0.5)
- Little to No Correlation (0.0 – 0.2)

When two independent variables are highly correlated, they are typically not both included in a model because they describe the same aspect and provide redundant information to the model. However, it should be noted that these correlation statistics were calculated on the variable as a set, whereas, for the categorical variables, the models only look at individual values (as per dummy coding described in Section 3.3). So, for example, although class and equivalent ESALs are strongly correlated as variables, their individual values (e.g. Hwy Class B and traffic group 3) may not be, and may both be valuable in a model.

4.4 Step 2b – Choose Model Type

Because it was not known which model would best suit the City's needs, three regression models were developed: a linear model, non-linear (exponential) model, and an ordinal logistic model. All of these models are empirical (based solely on regression analysis) because detailed measurements that would be necessary for a mechanistic-empirical model were not available. Soft computing methods were not considered as they require additional software and do not fit into the current framework.

Despite the fact that deterioration was measured on a discrete scale, models that required an interval dependent variable were still considered. The greater the number of states in a discrete scale, the more closely this measure resembles an interval variable. In this case, the discrete variable can take on up to 21 states, and a model with a dependent variable measured on an interval scale can produce reasonable results.

Regression analysis models the relationship between independent variables (e.g. surface material, traffic, etc.) and a dependent variable – in this case, structural adequacy score. Multiple linear regression and exponential regression are both deterministic models – they provide a single output for a given set of inputs. Logistic regression, on the other hand, takes into account the fact that a set of inputs may not always result in the same output. The output, in this case, is the probability that a set of inputs produces a particular output.

The linear model is the simplest to create, and was used as a starting point. Since, at first glance, the fit was relatively good, it is also used as a point of comparison to the other models. The exponential model was developed because pavement deterioration is expected to take this form – a relatively slow rate of deterioration that increases as the pavement nears the

end of its useful life. The ordinal logistic model was developed because it fits the form of the dependent variable (ordinal) and because it can be easily incorporated into a risk model since its output is in the form of a probability.

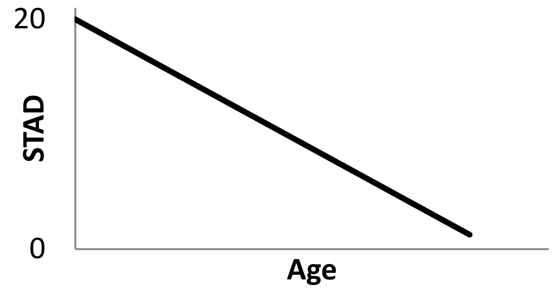
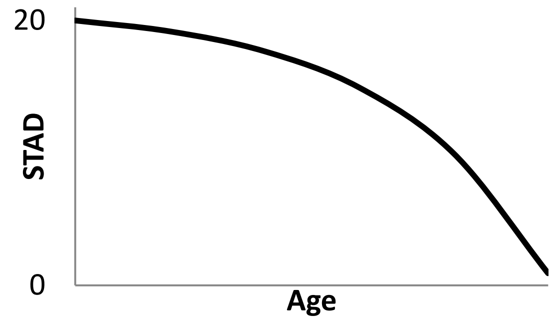
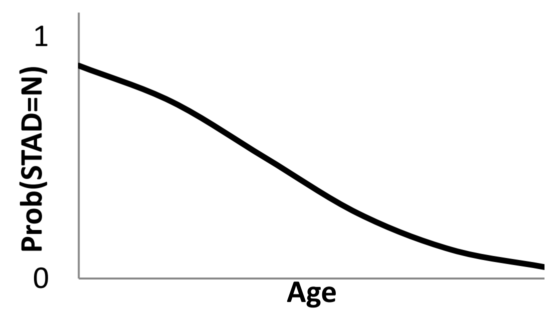
Because the distances between values in the dependent variable are unknown in ordinal logistic regression, fewer categories gives the potential for a more accurate model. This is because there are more data points for each combination of variables. For this reason, structural adequacy scores and age variables have been binned into as few values as possible while still including all the information necessary for decision-making. Details on these variables can be found in Section 4.3 Step 2a – Data Mining.

The original dataset is made up of a large majority of data points in structural adequacy category 7. This is an “unbalanced” dataset (more values in one category of the dependent variable than others) and skews the results of logistic regression. To solve this problem, a random selection of 13 records per structural adequacy category was used to create the model.

4.5 Step 3 – Develop Model Form

Table 13 provides a brief overview of the model forms. Parameter A, or the y-intercept, is set equal to 20 for the linear regression model and 21 for the exponential regression model. This reflects the fact that newly constructed roads with HCB, ICB or A/C surface types should have a structural adequacy score of 20 at age 0. Thus, the other independent variables discussed in Section 4.3 may only influence the slope, or rate, of deterioration.

Table 13 Overview of model types

Name	Base Equation	Sketch
Multiple Linear Regression	$STAD = A - BX - (CX \times age)$	
Exponential Regression	$STAD = A - BX - \mathrm{e}^{((C+DX)\times age)}$	
Ordinal Logistic Regression	$Prob(STAD_{bin} = N) = \dfrac{1}{(1 + e^{-\left(A+BX+((C+DX)\times age)\right)})}$	

4.6 Step 4 – Model Development

All of the regression models were developed using IBM SPSS Statistics software. Generally, the models were developed by starting with a simple form of the equation using age as the only independent variable. Other variables were added one by one to the model if they were found to significantly improve the results of the model. The optimal values of the parameters, the coefficients associated with the independent variables, were found using statistical software. In the case of multiple linear regression and exponential regression, the parameter values were found such that the sum of the squared difference of the expected and predicted structural adequacy scores was minimized. In the case of ordinal logistic regression, the maximum likelihood method was used – the parameters were optimized such that they provided the maximum likelihood that the data points occurred.

4.6.1 Multiple Linear Regression

The resulting best-fit linear regression equation is:

$$\begin{aligned} \mathit{Structural\ Adequacy\ Score} = 20 - (0.619\ \mathit{if\ resurfaced}, 0\ \mathit{if\ new\ construction}) - \\ (6.583\ \mathit{if\ surface\ type\ LCB\ or\ PRI}, 0\ \mathit{if\ HCB, ICB\ or\ AC}) - \\ ((0.18 + (0.049\ \mathit{if\ in\ traffic\ group}\ 3, 0\ \mathit{if\ in\ group}\ 1\ \mathit{or}\ 2) + \\ (0.072\ \mathit{if\ in\ Hwy\ Class\ B}, 0\ \mathit{ifnot}) + \\ (0.208\ \mathit{if\ in\ a\ rural\ environment}, 0\ \mathit{if\ not})) * \mathit{age}) \end{aligned} \quad [14]$$

Figure 14 shows the recommended linear model along with the structural adequacy data from the City of Oshawa. The upper and lower bounds of the model are found by inputting the set of independent variables that result in the highest and lowest structural adequacy scores respectively. The data points corresponding to these variable sets are also shown. Those data points in the "other" category do not belong to either the upper or lower datasets.

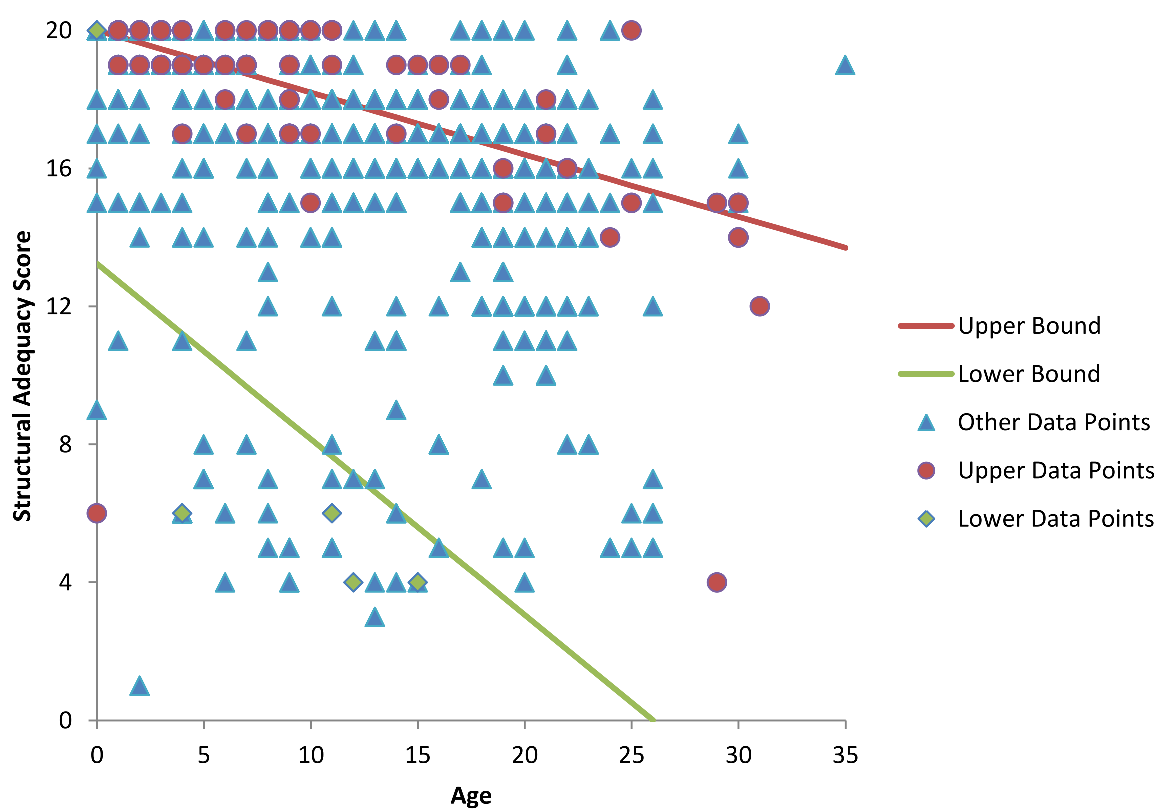


Figure 14 Multiple linear regression model

4.6.2 Exponential Regression

The exponential regression equation is:

$$\begin{aligned} Structural\ Adequacy\ Score &= 21 - (1.282\ if\ resurfaced, 0\ if\ new\ construction) - \\ &(8.277\ if\ LCB\ or\ PRI, 0\ if\ HCB, ICB\ or\ AC) - e^{0.072 \times age} \end{aligned} \quad [15]$$

Figure 15 shows the recommended linear model along with the structural adequacy data from the City of Oshawa. The upper and lower bounds of the model are shown and the curve corresponding to most of the data points (resurfaced and HCB, ICB or AC surface type), along with the data points corresponding to these lines.

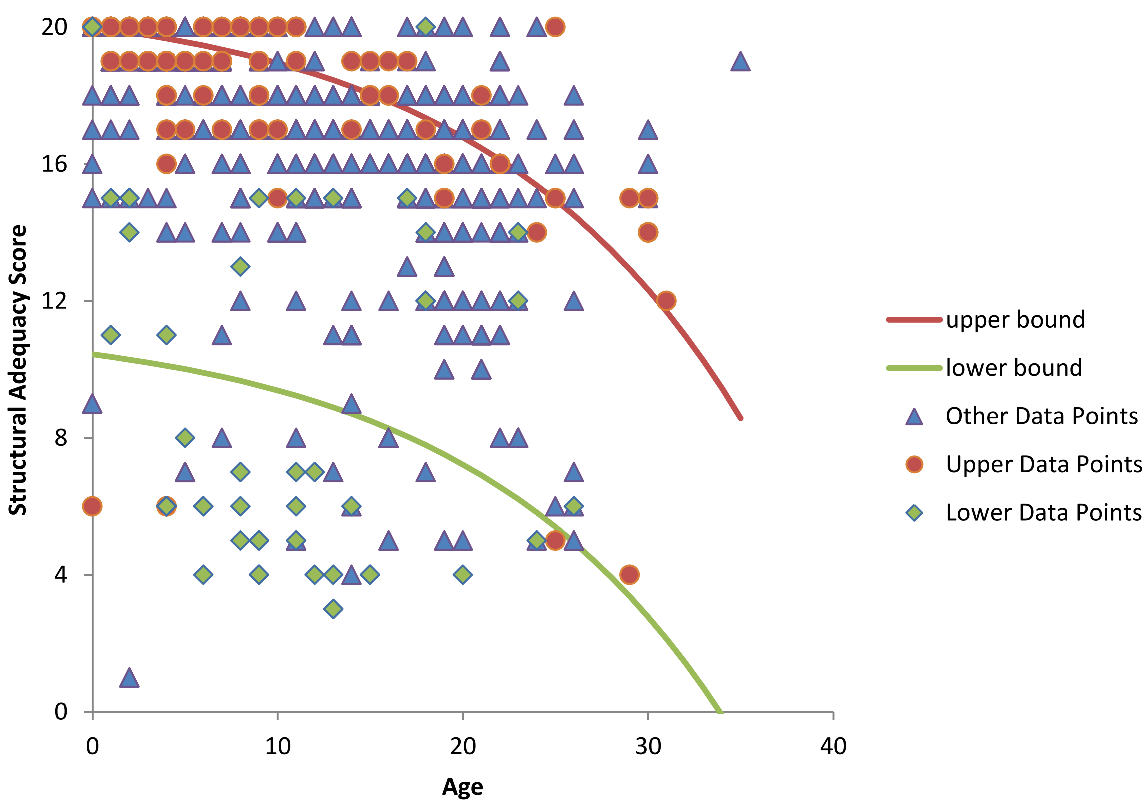


Figure 15 Exponential regression model

It should be noted that several aspects of this model were specified rather than derived from the data. The y-intercept, or the structural adequacy score at time 0, was set to 20 for newly constructed roads with a HCB, ICB or A/C surface type. Thus, many of the variables

described in section 2 could only influence the slope, or rate of deterioration. However, it was found that there was not a significant improvement in the Pearson's R^2 value by adding these terms.

4.6.3 Ordinal Logistic Regression

As was stated in Section 4.4, to eliminate bias, this model was developed using a "balanced" portion of the full dataset. The equations for this model can be found below.

$$Prob(STAD\_bin = 1) = \frac{1}{\left(1 + e^{\left(0.296-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1 or 2) = \frac{1}{\left(1 + e^{\left(-0.730-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1,2 or 3) = \frac{1}{\left(1 + e^{\left(-1.400-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1,2,3 or 4) = \frac{1}{\left(1 + e^{\left(-2.015-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1,2,3,4 or 5) = \frac{1}{\left(1 + e^{\left(-2.688-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1,2,3,4,5 or 6) = \frac{1}{\left(1 + e^{\left(-3.641-(0.372*age\_group)+(3.070\ if\ HCB,ICB\ or\ AC,0\ if\ LCB\ or\ PRI)\right)}\right)}$$

$$Prob(STAD\_bin = 1,2,3,4,5,6 or 7) = 1 \qquad [16]$$

Figure 16 shows the probability for a pavement section with a HCB, ICB or A/C surface type for each structural adequacy category over the age of the pavement.

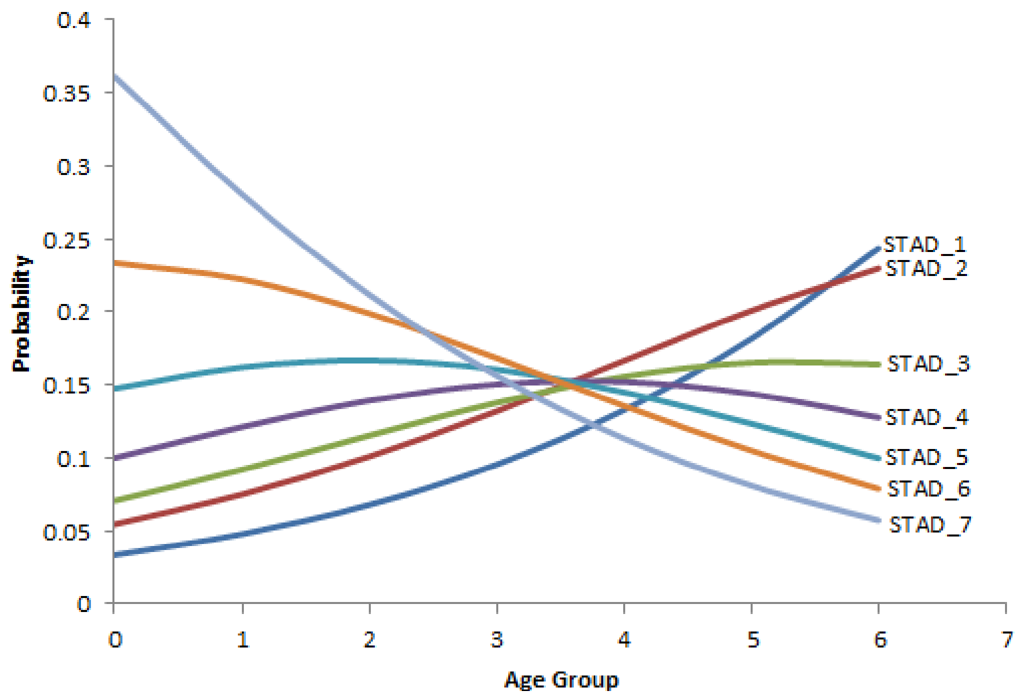


Figure 16 Ordinal logistic regression model

In this model, age has been grouped (see section 4.3.2.3) in 5 year increments. The structural adequacy scores have also been grouped (see section 4.3.1) into as few categories as possible while still retaining their relevance in the overall asset management plan. In approximately the first 15 years (up to age group 3) after it has been resurfaced, the pavement is most likely to be in Structural Adequacy group 7 (structural adequacy score 15 – 20), or good condition. Through ages approximately 15 to 20 (age groups 3 to 4), the probability of a pavement being in one structural adequacy category versus another is similar. In the later years (age 20/ age group 4 onwards) of its life, a pavement is likely to be in category 1 or 2, or poor condition.

4.7 Step 5 – Model Evaluation

The methods for evaluating a deterioration model vary with the model type. However, common evaluative measures include an evaluation of:

- Reasonable and significant parameter values,

- Plot of residuals,
- R^2, and
- Any assumptions.

Further detail about these measures can be found in Section 3.7.

4.7.1 Multiple Linear Regression Model Evaluation

Table 14 provides details on the parameter estimates.

Table 14 Multiple linear regression parameter estimates

Parameter	Estimated Parameter Value	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Resurfaced/New	0.619	0.105	0.412	0.826
Age	0.18	0.008	0.165	0.195
Surface Type	6.583	0.376	5.844	7.321
Traffic_group_3	0.049	0.013	0.024	0.074
Hwy_Class_B	0.072	0.018	0.037	0.107
Rural Environment	0.208	0.037	0.135	0.281

Because all of the confidence intervals do not cross 0, all of the parameter values are significant.

Error! Reference source not found. shows the residual values over the range of actual structural adequacy scores.

As the actual structural adequacy score decreases, the absolute values residual values become greater. This means that the model is able to predict data points that are in better condition more accurately. This makes sense, given that condition decreases with time, and most road sections are in relatively good condition at time 0. At higher structural adequacy scores (approximately structural adequacy score 18 and above), the model is more likely to overestimate condition, whereas at lower structural adequacy scores (approximately structural adequacy score 11 and below), the model is more likely to underestimate condition.

The R^2 value (calculated as 1 – (residual sum of squares)/(corrected sum of squares)) for this model is 0.50.

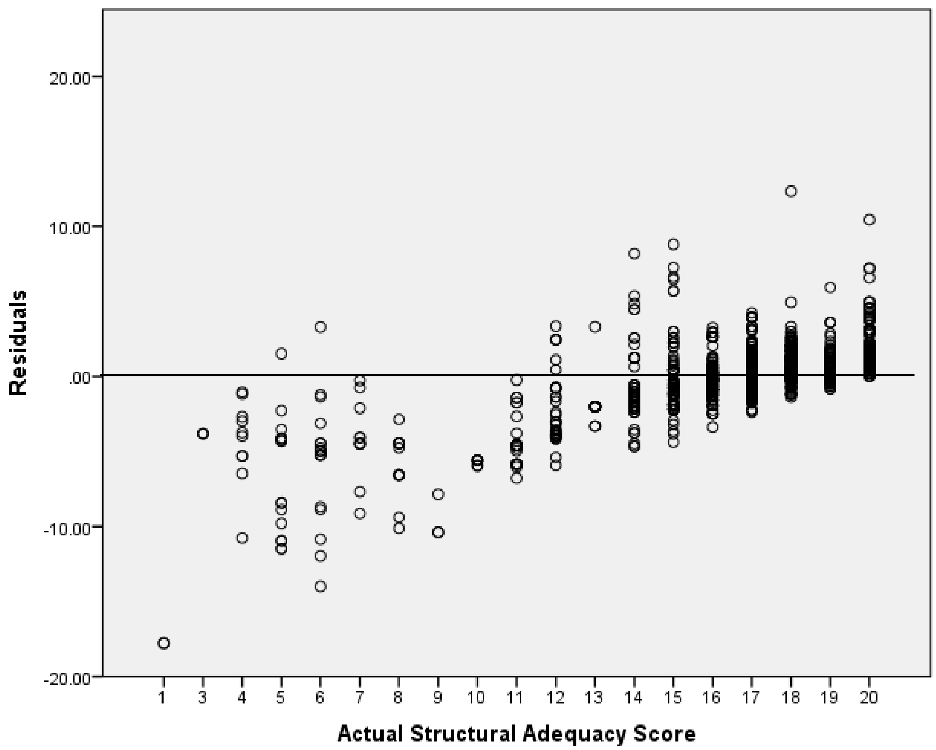


Figure 17 Multiple linear regression residuals

4.7.2 Exponential Regression Model Evaluation

Table 15 shows details of the parameter estimates.

Table 15 Exponential regression parameter estimates

Parameter	Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Resurfaced/New	1.282	0.084	1.117	1.447
Surface Type	8.277	0.277	7.734	8.819
Age	0.072	0.001	0.069	0.075

All the parameters are significant because the confidence intervals do not cross 0.

Similar to the linear regression model, two methods were used to evaluate the exponential model. The first is a plot of the residuals over their actual structural adequacy score as seen in Figure 18.

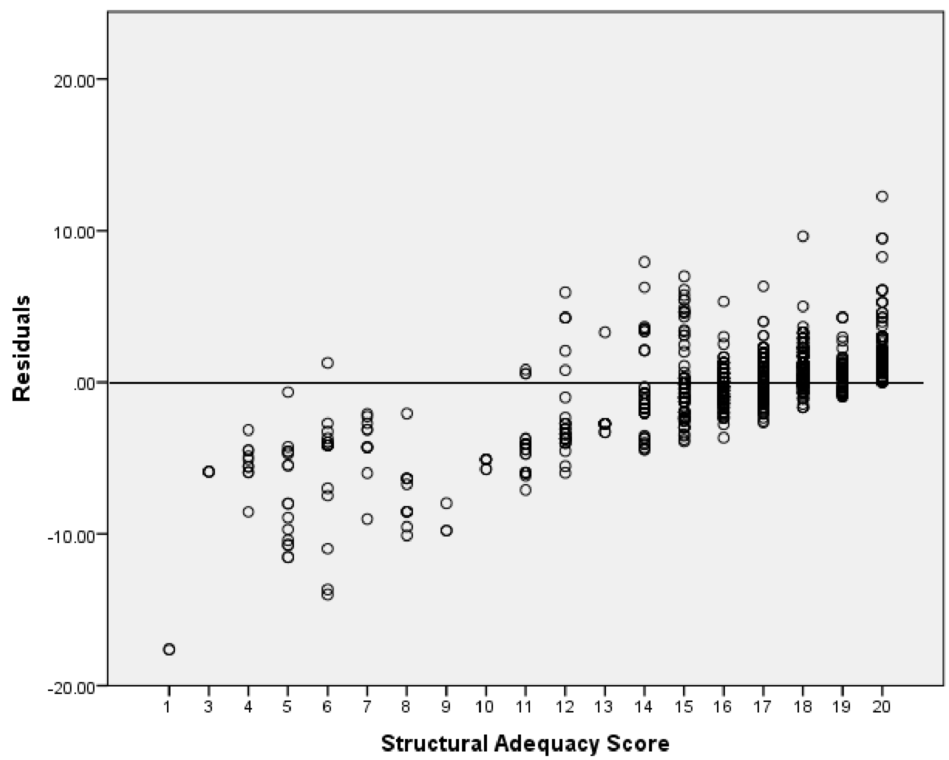


Figure 18 Exponential regression residuals

This plot is very similar to **Error! Reference source not found.**, a plot of the linear regression residuals. From this chart it can be seen that the absolute value of the residuals increases as the actual structural adequacy score decreases. Also, as the actual scores decrease, the predicted scores are more likely to be an overestimate of the actual score. This means that, similar to the linear model, the exponential model predicts the structural adequacy of those sections that actually have a higher structural adequacy score better than those that actually have a lower score. This is because condition decreases with time, and most road sections are in relatively good condition at time 0. At higher structural adequacy scores (approximately structural adequacy score 18 and above), the model is more likely to

overestimate condition, whereas at lower structural adequacy scores (approximately structural adequacy score 12 and below), the model is more likely to underestimate condition.

R^2 is calculated as 1 – (residual sum of squares)/(corrected sum of squares). For this model, the R^2 value is 0.45.

4.7.3 Ordinal Logistic Model Evaluation

The details of the parameter values for this model can be found in Table 16.

Table 16 Ordinal logistic regression parameter estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	2009_NEW_STAD_binned = 1	-0.296	0.626	0.223	1	0.637	-1.522	0.931
	2009_NEW_STAD_binned = 2	0.730	0.643	1.286	1	0.257	-0.531	1.991
	2009_NEW_STAD_binned = 3	1.400	0.656	4.548	1	0.033	0.113	2.686
	2009_NEW_STAD_binned = 4	2.015	0.666	9.153	1	0.002	0.710	3.320
	2009_NEW_STAD_binned = 5	2.688	0.677	15.78	1	0.000	1.362	4.015
	2009_NEW_STAD_binned = 6	3.641	0.703	26.79	1	0.000	2.262	5.020
Location	SURF_MAT = HCB, ICB, A/C	3.070	0.720	18.204	1	0.000	1.660	4.481
	SURF_MAT = LCB, PRI	0[a]	.	.	0	.	.	.
	AGE_YRLASTWK_2009_binned	-0.372	0.164	5.143	1	0.023	-0.693	-0.050

Link function: Logit.

a. This parameter is set to zero because it is redundant.

All the parameters except the intercept values for the two lowest structural adequacy categories are significant. Since the structural adequacy score variable has already been separated into as few categories as possible, the model will be evaluated despite this.

Figure 19 shows the frequencies of each structural adequacy category based on expected values from the model.

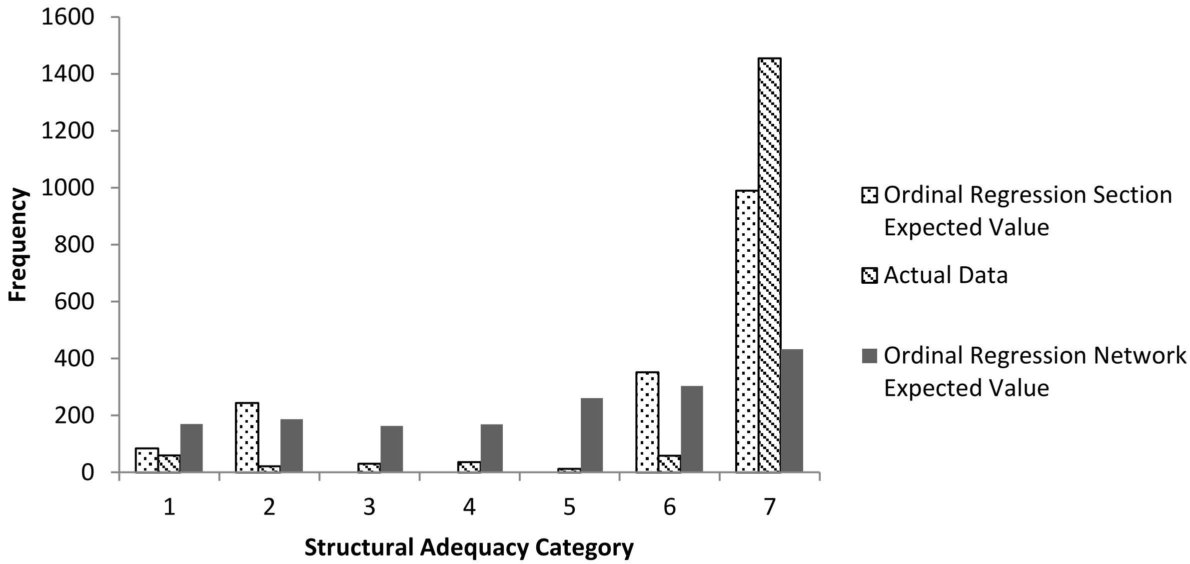


Figure 19 Ordinal logistic regression results

The section expected value is the structural adequacy category that the road section is most likely to be in – the structural adequacy category that has the highest probability according to the model. The model only predicts individual road section structural adequacy scores in four categories: 1, 2, 6, and 7. This prediction is shown to be incorrect when compared to the actual data. The network expected values – calculated by summing the probabilities of each structural adequacy category over the network – show that the model is also inaccurate in predicting the overall network condition.

When creating an ordinal regression model, proportional odds are assumed. The test of parallel lines, as shown in Table 17, checks this assumption.

Table 17 Test of parallelism

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	139.651			
General	115.177			
Difference		24.474	10	.006

Given that the significance of the difference of the -2 log likelihood values is low, in this case, the test of parallel lines does not support this assumption. This is likely because of the sparse and variable data at intermediate structural adequacy score values. A violation of this assumption also means that ordinal logistic regression may not be the appropriate model for this data.

There are several measures used to evaluate how well an ordinal regression model fits the data. One that is commonly used is the Pearson goodness-of-fit statistic and its associated deviance measure. In this case, these measures are not effective in evaluating the model because there are many possible combinations of variables that have very low, or 0, expected frequencies. This is primarily caused by the relatively large number of categories in both the dependent variable, and the age binned variable. In both of these cases though, it is impossible to reduce the number of categories without losing information that is relevant to the final use of the model. Therefore, the Pearson goodness-of-fit statistic and the deviance measure have not been included in this report.

The overall model, however, is significant. As can be seen in Table 18, the model with predictors is significantly better than the model without predictors.

Table 18 Model fitting information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	155.981			
Final	139.651			
Difference		16.330	2	.000

The pseudo-R^2 measures can be found in Table 19 for the ordinal logistic regression model.

Table 19 Pseudo-R^2 measures

Cox and Snell	.164
Nagelkerke	.168
McFadden	.046

The pseudo-R^2 statistics are all small for this model. This reinforces what has been discussed above – that the ordinal logistic regression model does not fit the data well.

One difficulty that greatly impacted the ordinal regression model in particular is that a number of possible combinations don't exist in the data set. This was the reason age was binned into ranges, but even with the grouping, there are many possible combinations (in fact, over 50%) that had no data. More data, particularly in the intermediate condition groups (structural adequacy 10-15), would likely lead to a better model.

4.8 Step 6 – Test Model

In terms of predicting the data that was used to create the models, multiple linear regression gave the best results with the highest R^2 value of any of the models. It should be noted that the pseudo-R^2 values found for the ordinal logistic regression cannot be directly compared to the R^2 values found for the linear and exponential regression models due to a difference in the calculation method and in the data sets used to create the models.

It should not be assumed that the model with the highest R^2 value is the "best" one. R^2 is very dependent on the data set used to create the model. More independent variables lead to a higher R^2 value. However, more independent variables also make the model more complicated and may result in over-fitting[1]. So while the linear model has the highest R^2 value, it must also be considered that it has the most independent variables.

R^2 also puts a relatively low importance on intermediate scores (structural adequacy scores less than 15) because there are fewer data-points at these ages. Thus, R^2 primarily reflects the fit of the model where data-points are plentiful – at higher structural adequacy scores.

One can also define the "best" model as the one with the most practical application. The primary use for this model is to work with the City of Oshawa's asset management system. Therefore, the prediction of road sections in intermediate and poor conditions is extremely important. The most important independent variable in these models is the time since the last work on the section was completed. However, at approximately 25 years, the standard error

[1] Over-fitting occurs when the model describes individual datapoints (including their errors) instead of an overall trend. The datapoints are "memorized" rather than generalized to form a trend.

(calculated as standard deviation over the square root of the number of observations) for the structural adequacy scores increases dramatically. Figure 20 shows the standard error of the structural adequacy scores over time.

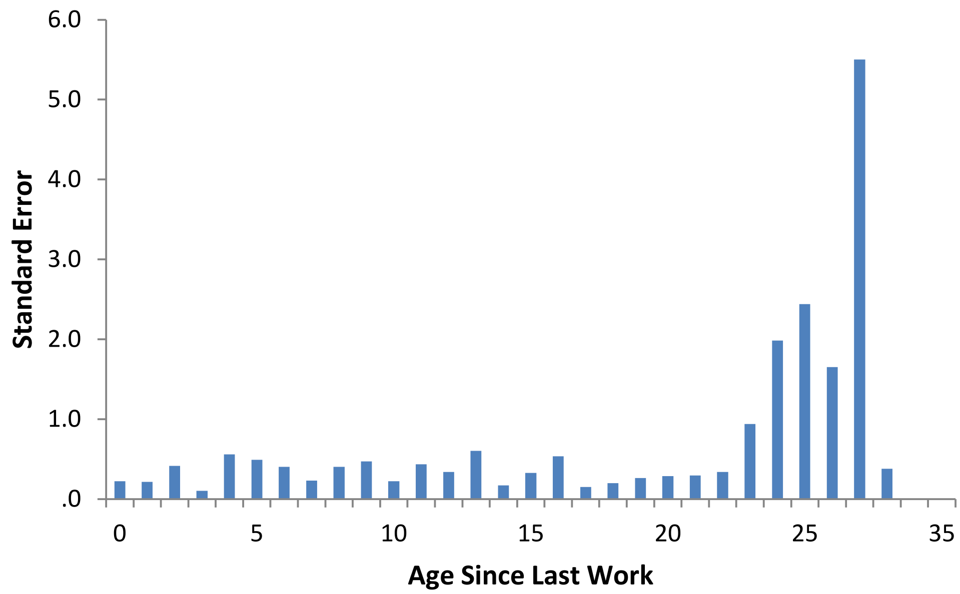


Figure 20 Standard error by age (based on year of last work)

Because the data is so variable at later ages, the structural adequacy scores are more difficult to accurately predict. The linear and exponential models are generally extrapolating at this point – their parameter values having been greatly influenced by the more plentiful and less variable data found at earlier ages. The ordinal regression model though, depends on the data at these later ages (and corresponding lower condition scores) and does not perform as well.

Another factor that must be considered when evaluating the models is how reasonable the results are when compared to what is expected. Because the models are extrapolating at higher ages, it is particularly important to check that the predicted values make sense at these ages. The rate of deterioration is typically expected to increase with age. The exponential model shows this increase in the rate of deterioration, where the linear model does not.

Although in a theoretical sense the ordinal regression model should provide the best results because it assumes that that the dependent variable is ordinal (as opposed to linear, which is assumed for the deterministic models) and because its output are in the form of a probability,

which can easily be incorporated into a risk-based system, it does not provide the best results. The ordinal regression model does not accurately predict structural adequacy scores in the intermediate range. Since this range is extremely important in the overall asset management program, it is not recommended that this model be used.

Because the exponential model provides reasonable results with relatively few independent variables, it is recommended as the best choice for the City of Oshawa's asset management plan.

4.9 Summary

This chapter presents an application of the framework outlined in Chapter 3, applied to a dataset from the City of Oshawa. Because it was not known which would best suit the data, three model types were selected: multiple linear regression, exponential regression and ordinal logistic regression. Approximately 1700 records were used to create the linear and exponential models. To create the ordinal logistic model, a smaller set of around 90 records was used to reduce bias.

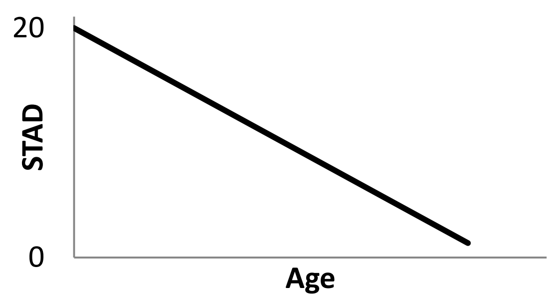
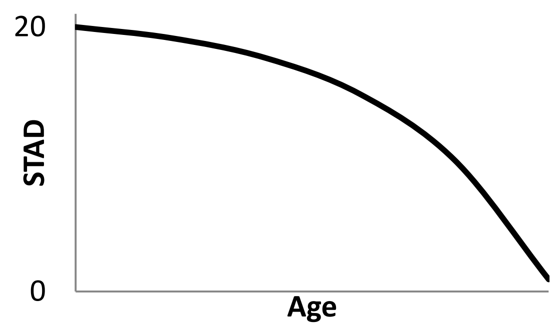
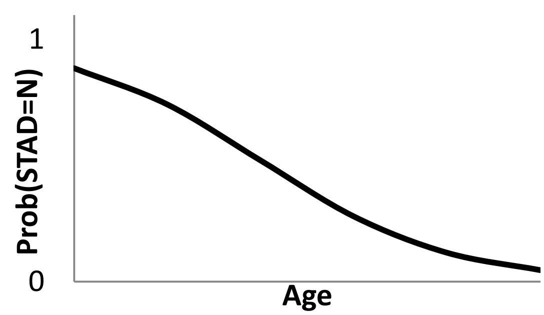
A summary of the pavement deterioration models developed for the City of Oshawa can be found in Table 20.

The parameter values were found to be reasonable and significant for both the linear and exponential models. However, some parameters are insignificant in the ordinal logistic model. The R^2 values are 0.45 and 0.5 for the exponential and linear models respectively, and are quite low (ranging from 0.05 to 0.17 depending on the measure) for the ordinal logistic model. Based on an evaluation of the models, both the linear and exponential models provide reasonable results, but the ordinal logistic model does not.

The linear and exponential models are generally extrapolating to predict intermediate and low condition states. These predictions are extremely important in the overall asset management system because predicted condition triggers particular maintenance or rehabilitation treatments. The City of Oshawa expects that a pavement's rate of deterioration will increase with time. The exponential model shows an increase in the rate of deterioration with time. Because the exponential model provides reasonable results with relatively few independent

variables, it is recommended as the best choice for the City of Oshawa's asset management plan.

Table 20 Summary of City of Oshawa models

Name	Dependent Variable	Independent Variables	Fit	Sketch
Multiple Linear Regression	Structural Adequacy Score	Age New/Resurfaced Surface Type Traffic (Equivalent ESALs) Class Environment	$R^2 = 0.50$	
Exponential Regression	Structural Adequacy Score	Age New/Resurfaced Surface Type	R^2=0.45	
Ordinal Logistic Regresssion	Probability of Structural Adequacy Score Bin	Age Surface Type	Nagelkerke $R^2 = 0.17$	

5 Case Study: Trunk Sewer Deterioration Model

The factors that affect sewer deterioration have not been agreed upon in literature (see Section 2.2). The purpose of this case study is to determine if the condition of large diameter trunk sewer pipes can be predicted by their age and material.

5.1 Step 1a – Compile Data

The dataset used for this study was provided by a sewer condition assessment firm. The data consists of a sample of large-diameter sewer pipes from a Canadian municipality. The final dataset was assembled using Microsoft Access and contains 1315 records (although only around 200 will be used in the final model). Each record represents sewer pipes, ranging in length from 2m to 550m.

5.2 Step 1b – Research Model Types

Sewer deterioration is most often modelled with a probabilistic or soft computing type model. A brief overview of the research that has been performed, and the models that apply to sewer deterioration, can be found in Chapter 2.

5.3 Step 2a – Data Mining

The following section provides an overview of the variables used in the sewer deterioration model.

5.3.1 Dependent Variable – Structural Condition Grade

Sewer deterioration is commonly separated into two general categories: structural and service. Structural defects include cracks, corrosion, openings in joints, breaks and holes; these defects decrease the sewer's structural capacity. Structural deterioration can occur through several mechanisms, including four-point fracture, subsidence, and fabric decay. Service defects include infiltration, root intrusion, encrustation and debris; by gradually reducing cross-sectional area, these defects reduce the sewer's hydraulic capacity. In this case, the structural grade of the sewer section is the dependent variable.

The structural grade has been determined using the Water Research Council's (WRc) grading system. Grades are assigned based on the deficiencies noted in a CCTV inspection of the

sewer. Grades are integer values, and range from 1 (good condition) to 5 (collapsed). Figure 21 shows the distribution of grades for the entire dataset.

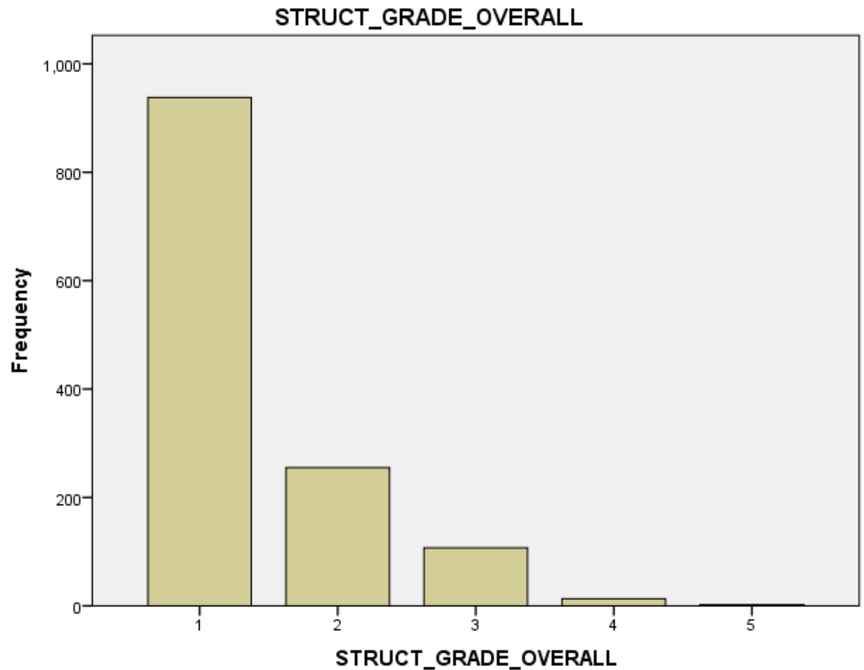


Figure 21 Condition grade histogram

Most of the sewer sections are in good condition, with progressively fewer sections in each of the lower grades. Because there are relatively few sewer sections in the lower grades, the scores have been grouped together into poor condition (grades 3,4 and 5) and good condition (grades 1 and 2). Then, to provide an unbiased dataset, around 100 records from each of the poor and good condition sets (~200 records total) were used in the analysis.

5.3.2 Independent Variables

Independent variables are considered inputs to the model. They help to predict the condition grade.

5.3.2.1 Age

Age is measured from the date of original construction to the year of inspection. Inspections on the network took place from 1996 to 2011. Figure 22 shows the distribution across the dataset.

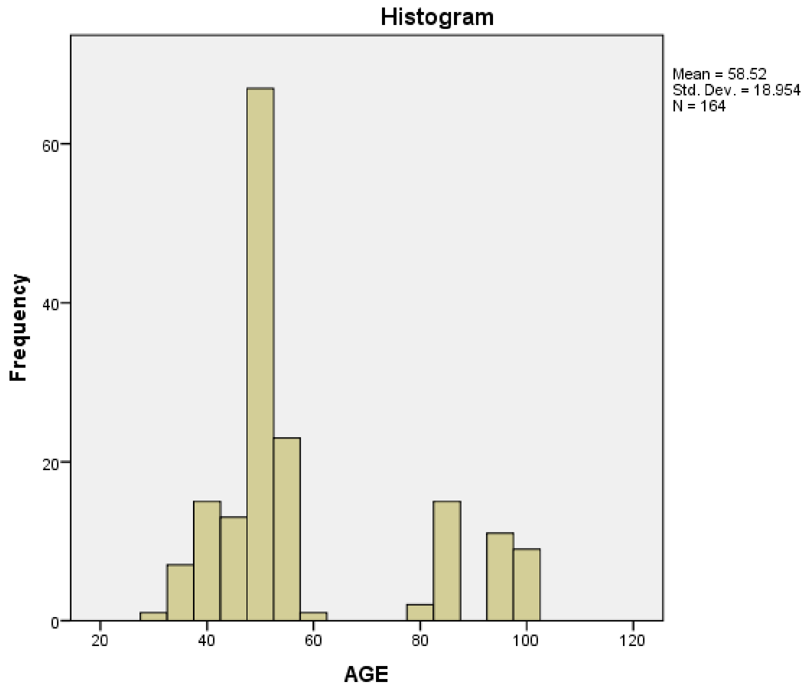


Figure 22 Age histogram

The sewer sections range in age from 30 to 99 years, with most sections built around 50 years before their inspection dates.

5.3.2.2 Material

Eighty-six percent of the sample set is made up of concrete pipe. The remaining 14% pipe segments are made of brick.

5.3.3 Correlations

Table 21 presents the correlation between variables.

Table 21 Variable correlations

	Structural Grade	Age	Material
Age	No correlation		
Material	Insignificant correlation	eta: 0.91	

Neither age nor material has a significant correlation with structural grade. However, there is a very high correlation between age and material as material choices changed with construction techniques over the past 100 years – brick was a popular material choice for trunk sewers in the early 20th century, but is no longer used in sewer construction.

5.4 Step 2b – Choose Model Type

A logistic regression model will be used in this case. This method was chosen because the dependent variable is ordinal, and because sewer deterioration is thought to occur (for the most part) in discrete steps, rather than gradually. Also, both Younis et al.(2010a) and Ana et al. (2009) used logistic regression models to investigate the factors that affect sewer deterioration.

5.5 Step 3 – Develop Model Form

The model form is shown in the following equation.

$$\log\left(\frac{P(poor\ condition)}{P(good\ condition)}\right) =$$

$$B_0 + B_1(1\ if\ material\ is\ concrete, 0\ if\ it\ is\ brick) + B_2(age) \qquad [17]$$

Because the purpose of this model is to determine if material and age influence deterioration, both independent variables have been entered into the model despite their very high correlation. It is expected that one or, most likely both, variables will not be significant in the final model.

5.6 Step 4 – Model Development

The binary logistic model is found below.

$$\log\left(\frac{P(poor\ condition)}{P(good\ condition)}\right) =$$

$$0.602 - 0.14(age) + (0.904\ if\ brick, 0\ if\ concrete) \qquad [18]$$

Figure 23 shows the probability for a sewer section to be in good condition at a particular age.

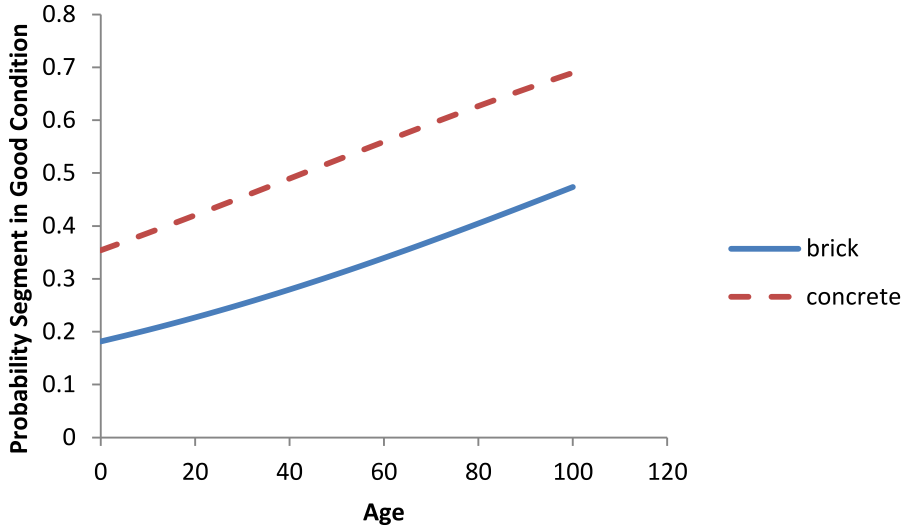


Figure 23 Binary logistic regression model

This model (incorrectly) shows that the probability that a section is in good condition increases with age. Concrete pipe sections are more likely to be in good condition than brick sewers.

5.7 Step 5 – Model Evaluation

The details of the parameter values for this model can be found in Table 22.

Table 22 Binary logistic regression parameter estimates

Parameter Values for Dependent Variable "Poor Condition"	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	0.602	0.714	0.712	1	0.399			
AGE	-0.014	0.013	1.113	1	0.292	0.986	0.96	1.012
[MATERIAL=Brick]	0.904	0.674	1.798	1	0.18	2.47	0.659	9.265
[MATERIAL=Concrete]	0[b]	.	.	0	.	.	.	.

a. The reference category is: Good Condition.

b. This parameter is set to zero because it is redundant.

All of the parameters have low Wald scores and are insignificant. The parameter value for age is negative, when it is expected to be positive, since the probability that a section is in poor condition should increase with age.

The Cox and Snell, Nagelkerke, and McFadden pseudo R^2 values are also extremely low – all have values below 0.02. The model fit statistics, shown in Table 23, show that the model with independent variables (material and age) included is not significantly better than a model without.

Table 23 Model fitting information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	96.779			
Final	94.911			
Difference		1.868	2	.393

These statistics show that the model, with material and age as independent variables, does not predict sewer deterioration at all well. In fact, the model shows that condition a sewer is more likely to be in good condition when it is older.

According to Figure 4, since the model evaluation results are not acceptable, there are two options for the modeller. These are based on the question of whether the model type is still appropriate. As was mentioned in Section 5.2, other researchers have used this model type successfully. For this dataset, however, it may not be appropriate.

If the model type is appropriate, other variables are likely to influence deterioration. These might include: depth, location, size, construction technique, contractor, inspector, construction specifications, soil conditions, water table depth relative to sewer depth, etc. Davis et al. (2001a) and Ana et al. (2009) provide overviews of the factors that are likely to affect sewer deterioration.

Alternatively, the same variables might be used in another model type. A Markov model or method using soft computing techniques may provide more reasonable results.

5.8 Summary

In this chapter, the framework described in Chapter 3 is applied to create a trunk sewer deterioration model. Age, material and condition data were obtained for a Canadian municipality, and applied in an logistic regression model.

The model did not fit the data well. The model with variables was not found to fit the data significantly better than the model without age and material included. Thus, in this particular case, age and material do not prove to affect condition.

To obtain a better model, other variables (such as construction technique, soils information, etc.) might be used; or alternatively, a different model type (such as a Markov model) might be applied.

6 Conclusions and Recommendations

Deterioration modelling is an integral part of infrastructure asset management. Models are used to predict future condition and plan maintenance and rehabilitation treatments. The purpose of this research was to develop a framework to create infrastructure deterioration models – particularly for roads and sewers.

This paper presents a brief literature review outlining the deterioration process and the factors that might affect it. An overview of the types of deterioration models is also included, presenting the advantages and disadvantages of each. Existing deterioration model frameworks have also been examined. A deterioration modelling framework was then proposed. This includes information on infrastructure data, choosing a deterioration model, and how a model might be developed, evaluated and tested.

The framework was then applied in two case studies. The first is a comparison of three pavement deterioration models created for the City of Oshawa for use in their asset management system. Three models were developed and compared – a linear regression model, an exponential regression model, and an ordinal regression model. The ordinal logistic model did not produce good results, likely due to the fact that the data most relevant to this model found at intermediate condition states was variable and sparse. The linear and exponential models were generally extrapolating at the higher ages (25 and up) related to these intermediate condition states. It was found that the model with the highest Pearson's R^2 value (the linear model) was not necessarily the best suited to the City's needs. The City expected the rate of deterioration to increase as the pavement aged and so the exponential model was selected.

The second case study involved modelling sewer deterioration in large diameter trunk sewers. The factors that influence sewer deterioration are not agreed upon in literature. In this case, the relationship between age, material and condition was explored using a logistic regression model. It was found that, for this dataset, age and material do not significantly affect condition. Based on the proposed framework, this model could be redeveloped using other variables, or another model type, such as a Markov model or soft computing technique, may be applied.

To ensure its flexibility, the proposed deterioration framework should be applied to other model types (particularly soft computing) and to other types of infrastructure. Also, further research on how a model is to be tested (e.g. its impact on life cycle cost and treatment selection) should be investigated.

References

International infrastructure management manual. 2006, 3rd edn, Association of Local Government Engineering N.Z. Inc (INGENIUM), Thames, New Zealand.

Ahammed, M. & Tighe, S. 2008. Statistical Modeling in Pavement Management: Do the Models Make Sense? *Transportation Research Record: Journal of the Transportation Research Board*, 2084(-1): 3-10.

Al-Barqawi, H. & Zayed, T. 2008. Infrastructure Management: Integrated AHP/ANN Model to Evaluate Municipal Water Mains' Performance. *Journal of Infrastructure Systems*, 14(4): 305-318.

Ana, E., Bauwens, W., Pessemier, M., Thoeye, C., Smolders, S., Boonen, I. & De Gueldre, G. 2009. An investigation of the factors influencing sewer structural deterioration. *Urban Water Journal*, 6(4): 303-312.

Ana, E.V. & Bauwens, W. 2010. Modeling the structural deterioration of urban drainage pipes: the state-of-the-art in statistical methods. *Urban Water Journal*, 7(1): 47-59.

Ariaratnam, S.T., El-Assaly, A. & Yang, Y. 2001. Assessment of Infrastructure Inspection Needs Using Logistic Models. *Journal of Infrastructure Systems*, 7(4): 160.

Baik, H.S., Jeong, H.S. & Abraham, D.M. 2006. Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems. *Journal of Water Resources Planning and Management - ASCE*, 132(1): 15-24.

Baur, R. & Herz, R. 2002. Selective inspection planning with ageing forecast for sewer types. *Water science and technology : a journal of the International Association on Water Pollution Research*, 46(6-7): 389-396.

Black, M., Brint, A.T. & Brailsford, J.R. 2005. A Semi-Markov Approach for Modelling Asset Deterioration. *Journal of the Operational Research Society*, 56(11): 1241-1249.

Bubtiena, A.M., Elshafie, A.H. & Jafaar, O. 2011. Application of Artificial Neural networks in modeling water networks. IEEE, 50.

Canadian Strategic Highway Research Program (C-SHRP) 2000, *Pavement Design and Performance: Current Issues and Research Needs*, Canadian Strategic Highway Research Program (C-SHRP), Ottawa, ON.

Cattan, J. & Mohammadi, J. 1997. Analysis of Bridge Condition Rating Data Using Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*, 12(6): 419-429.

Chang, J., Chen, S., Chen, D. & Liu, Y. 2008. Rutting Prediction Model Developed by Genetic Programming Method Through Full Scale Accelerated Pavement Testing. IEEE, 6: 326.

Chughtai, F. & Zayed, T. 2008. Infrastructure condition prediction models for sustainable sewer pipelines. *Journal of Performance of Constructed Facilities*, 22(5): 333-341.

Davies, J. 2001. The structural condition of rigid sewer pipes: a statistical investigation. *Urban Water,* 3(4): 277-286.

Davies, J.P., Clarke, B.A., Whiter, J.T. & Cunningham, R.J. 2001a. Factors influencing the structural deterioration and collapse of rigid sewer pipes. *Urban Water,* 3(1–2): 73-89.

Davies, J.P., Clarke, B.A., Whiter, J.T., Cunningham, R.J. & Leidi, A. 2001b. The structural condition of rigid sewer pipes: a statistical investigation. *Urban Water,* 3(4): 277-286.

Dirksen, J. & Clemens, F.H.L.R. 2008. Probabilistic modeling of sewer deterioration using inspection data. *Water science and technology : a journal of the International Association on Water Pollution Research,* 57(10): 1635-1641.

Eldin, N.N. & Senouci, A.B. 1995. A Pavement Condition-Rating Model Using Backpropagation Neural Networks. *Computer-Aided Civil and Infrastructure Engineering,* 10(6): 433-441.

Elhag, T.M.S. & Wang, Y. 2007, *Risk Assessment for Bridge Maintenance Projects: Neural Networks versus Regression Techniques,* ASCE.

Flintsch, G.W. & Chen, C. 2004. Soft Computing Applications in Infrastructure Management. *Journal of Infrastructure Systems*, 10(4): 157-166.

Fwa, T.F. & Chan, W.T. 1993. Priority Rating of Highway Maintenance Needs by Neural Networks. *Journal of Transportation Engineering,* 119(3): 419-432.

Haas, R. (ed) 1997, *Pavement design and management guide,* Transportation Association of Canada, Ottawa.

Henning, T.F.P., Costello, S.B. & Watson, T.G. 2006, *A review of the HDM/dTIMS pavement models based on calibration site data,* Land Transport New Zealand, Wellington, New Zealand.

Kerali, H.R., Robinson, R. & Paterson, W.D.O. 1998. Role of the New HDM-4 in Highway Management. *4th International Conference on Managing Pavements,* Transport Research Laboratory, Durban, South Africa, 801.

Kleiner, Y. & Rajani, B. 2001. Comprehensive review of structural deterioration of water mains: statistical models. *Urban Water,* 3(3): 131-150.

Kleiner, Y., Rajani, B. & Sadiq, R. 2006. Failure risk management of buried infrastructure using fuzzy-based techniques. *Journal of Water Supply Research and Technology - AQUA,* 55(2): 81-94.

Kleiner, Y. 2001. Scheduling Inspection and Renewal of Large Infrastructure Assets. *Journal of Infrastructure Systems,* 7(4): 136.

Konig, A. 2005, *CARE-S WP2 External corrosion model description*, SINTEF Technology and Society, Norway.

Le Gat, Y. 2008. Modelling the deterioration process of drainage pipelines. *Urban Water Journal,* 5(2): 97-106.

Lou, Z., Gunaratne, M., Lu, J.J. & Dietrich, B. 2001. Application of Neural Network Model to Forecast Short-Term Pavement Crack Condition: Florida Case Study. *Journal of Infrastructure Systems,* 7(4): 166.

Madanat, S., Mishalani, R. & Ibrahim, W.H.W. 1995. Estimation of Infrastructure Transition Probabilities from Condition Rating Data. *Journal of Infrastructure Systems,* 1(2): 120-125.

Micevski, T., Kuczera, G. & Coombes, P. 2002. Markov Model for storm water pipe deterioration. *Journal of Infrastructure Systems,* 8(2): 49.

Municipal Transportation Division, Ministry of Transportation 1991, *Inventory Manual for Municipal Roads*, Revised edn, The Research and Development Branch Ministry of Transportation Ontario, Downsview, Ontario.

Najafi, M. & Kulandaivel, G. 2005. Pipeline Condition Prediction Using Neural Network Models. ASCE, 180: 61.

National Guide for Sustainable Municipal Infrastructure (NGSMI). 2002. Decision Making and Investment Planning, Ottawa: NGSMI.

Olden, J.D. & Jackson, D.A. 2002. Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling,* 154(1–2): 135-150.

Osman, H. & Bainbridge, K. 2011. Comparison of statistical deterioration models for water distribution networks. *Journal of Performance of Constructed Facilities,* 25(3): 259.

Parkman, C., Hallet, J., Henning, T. & Trapper, M. 2003, *Pavement Deterioration Modelling in Long Term Performance Based Contracts: How Far Does it Mitigate the Risk for Client and Contractor?*, 21st ARRB and REAAA Conference, Cairns, New Zealand.

Rajani, B. & Kleiner, Y. 2001. Comprehensive review of structural deterioration of water mains: physically based models. *Urban Water*, 3(3): 151-164.

Raymond, C., Tighe, S., Haas, R. & Rothenburg, L. 2003. Development of Canadian asphalt pavement deterioration models to benchmark performance. *Canadian Journal of Civil Engineering*, 30(4): 637-643.

Saba, R.G. 2006, *Pavement Performance Prediction Models Project*, NordFoU, Norway.

Scheidegger, A., Hug, T., Rieckermann, J. & Maurer, M. 2011. Network condition simulator for benchmarking sewer deterioration models. *Water research*, 45(16): 4983-4994.

Schram, S.A. 2008, *Mechanistic-empirical modeling and reliability in network-level pavement management*, North Dakota State University.

Shekharan, A. 2000. Solution of Pavement Deterioration Equations by Genetic Algorithms. *Transportation Research Record: Journal of the Transportation Research Board*, 1699(-1): 101-106.

Syachrani, S., Jeong, H. & Chung, C. 2011. Dynamic Deterioration Models for Sewer Pipe Network. *Journal of Pipeline Systems Engineering and Practice*, 2(4): 123-131.

Tighe, S., He, Z. & Haas, R. 2001. Environmental Deterioration Model for Flexible Pavement Design: An Ontario Example. *Transportation Research Record: Journal of the Transportation Research Board*, 1755(-1): 81-89.

Tran, D., Ng, A., Perera, B., Burn, S. & Davis, P. 2006. Application of probabilistic neural networks in modelling structural deterioration of stormwater pipes. *Urban Water Journal*, 3(3): 175-184.

Tran, D.H., Ng, A.W.M. & Perera, B.J.C. 2007. Neural networks deterioration models for serviceability condition of buried stormwater pipes. *Engineering Applications of Artificial Intelligence*, 20(8): 1144-1151.

Tran, D.H., Perera, B. & Ng, A. 2009. Comparison of Structural Deterioration Models for Stormwater Drainage Pipes. *Computer Aided Civil and Infrastructure Engineering*, 24(2): 145-156.

Tran, H.D., Perera, B. & Ng, A. 2010. Markov and Neural Network Models for Prediction of Structural Deterioration of Storm-Water Pipe Assets. *Journal of Infrastructure Systems*, 16(2): 167-171.

Tran, H.D. 2007, *Investigation of deterioration models for stormwater pipe systems*, Victoria University. Architectural, Civil and Mechanical Engineering.

Transportation Association of Canada (TAC), Tighe, S. & Goodman, S. (eds) 2012, *Pavement Asset Design and Management Guide*, Transportation Association of Canada, Ottawa, Ontario.

Ullidtz, P. 1999. Deterioration Models for Managing Flexible Pavements. *Transportation Research Record: Journal of the Transportation Research Board,* 1655(-1): 31-34.

Vanier, D.J. 2001. Why industry needs asset management tools. *Journal of Computing in Civil Engineering,* 15(1): 35-43.

Wang, K.C.P. & Li, Q. 2011. Pavement Smoothness Prediction Based on Fuzzy and Gray Theories. *Computer-Aided Civil and Infrastructure Engineering,* 26(1): 69-76.

Wirahadikusumah, R., Abraham, D. & Iseley, T. 2001. Challenging Issues in Modeling Deterioration of Combined Sewers. *Journal of Infrastructure Systems,* 7(2): 77.

Wolters, A., McGovern, G. & Hoerner, T. 2006. Development of a Tool to Assess the Quality of Collected Pavement Management Data. *Transportation Research Record: Journal of the Transportation Research Board,* 1974(-1): 37-46.

Younis, R. & Knight, M.A. 2010a. A probability model for investigating the trend of structural deterioration of wastewater pipelines. *Tunnelling and Underground Space Technology,* 25(6): 670-680.

Younis, R. & Knight, M.A. 2010b. Continuation ratio model for the performance behavior of wastewater collection networks. *Tunnelling and Underground Space Technology incorporating Trenchless Technology Research,* 25(6): 660-669.

Yu, J. 2005, *Pavement service life estimation and condition prediction*, University of Toledo.